

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE ECONOMIA E RELAÇÕES INTERNACIONAIS
GRADUAÇÃO EM CIÊNCIAS ECONÔMICAS

RAFAEL TEODORO CUPERTINO

IMPACTOS DA INTELIGÊNCIA ARTIFICIAL NA ECONOMIA
MUNDIAL

UBERLÂNDIA – MG

2023

RAFAEL TEODORO CUPERTINO

**IMPACTOS DA INTELIGÊNCIA ARTIFICIAL NA ECONOMIA
MUNDIAL**

Monografia apresentada ao Instituto de Economia e Relações Internacionais da Universidade Federal de Uberlândia como requisito parcial para obtenção do título de bacharel em Ciências Econômicas.

Orientador: Professor Dr. Germano Mendes de Paula

Uberlândia – MG
27 de Junho de 2023

RAFAEL TEODORO CUPERTINO

IMPACTOS DA INTELIGÊNCIA ARTIFICIAL NA ECONOMIA MUNDIAL

Monografia apresentada ao Instituto de Economia e Relações Internacionais da Universidade Federal de Uberlândia como requisito parcial para obtenção do título de bacharel em Ciências Econômicas.

Orientador: Professor Dr. Germano Mendes de Paula

Uberlândia - MG, 27 de Junho de 2023.

BANCA EXAMINADORA

Professor Dr. Germano Mendes de Paula
Orientador

Professor Dr. Cássio Garcia Ribeiro Soares da Silva

Professor Dr. Marcelo Sartorio Loral

AGRADECIMENTOS

Vindo de uma mudança de curso de graduação, transcorri seis anos dentro dessa instituição, e agora, finalmente, posso alcançar a satisfação de concluir esse ciclo.

Sem o suporte e comprometimento dos meus pais, que desde a minha infância priorizaram a minha educação, não seria possível completar essa passagem. A vocês, Fernando e Idelte, eu dedico todo o meu amor e a minha gratidão. Suas dedicações foram recompensadas!

À minha irmã Ana, obrigado. Seu apoio foi imensurável e incessante. Essa é apenas uma das etapas na minha vida, e conto com essa enorme disposição que teve, proporcionada pela nossa amizade, para me ajudar a enfrentar as próximas.

Ao meu orientador, Germano, devo muito respeito e apreço. Não só pelo imenso auxílio e atenção durante a elaboração dessa monografia, mas também pela experiência de duas disciplinas ministradas com excelência. Os desafios foram grandes, mas mudaram minha percepção como discente e me ajudaram a melhorar como estudante.

Agradeço também aos docentes do Instituto de Economia e Relações Internacionais que fizeram parte da minha graduação, foram eles que reanimaram minha posição como aluno e me mostraram uma perspectiva mais humanista dessa área de estudo.

Por fim, aos amigos que estiveram presente durante esse processo árduo e me deram motivação, e pelas novas amizades que fiz nesse caminho, que me deram mais alegria e leveza para continuar nessa trajetória, muito obrigado!

RESUMO

A inteligência artificial (AI) é um dos assuntos mais importantes no atual contexto da Indústria 4.0, com potencial de transformar o cenário econômico e social de forma intensa e abrupta. Devido a sua capacidade de, dentro do seu limite vigente, reproduzir práticas humanas, essa tecnologia possui uma imensa área de difusão, permitindo que ela possa ser implementada em diversas áreas desde que seu uso seja preferencial, ou que não possa ser feito, por um humano. Nessa ótica, o objetivo dessa monografia é expor as possibilidades resultadas da IA, analisando de que forma ela interage com o setor econômico e social, e modifica as técnicas e percepções dos indivíduos nesse meio. Portanto, será investigado o nível de aplicação da IA, assim como os desafios éticos e morais que a envolvem junto à sua regulação, seu modelo atual de maior impacto – IA generativa – e por fim, seu latente grau disruptivo.

Palavras-chave: Inteligência artificial, Tecnologia da Informação, Ética na IA, Regulação da IA, IA generativa, ChatGPT, Trabalho.

ABSTRACT

Artificial intelligence is one of the most important issues in the current context of Industry 4.0, with the potential to transform the economic and social landscape intensely and abruptly. Due to its capacity to reproduce human practices within its current limitations, this technology has a vast scope of diffusion, allowing it to be implemented in various areas where its use is preferable or where it cannot be done by a human. From this perspective, the objective of this study is to present the possibilities resulting from AI, analyzing how it interacts with the economic and social sectors and modifies the techniques and perceptions of individuals within this context. Therefore, it will be investigated the level of AI application, as well as the ethical and moral challenges involved in it among its regulation, its current model of greatest impact - generative AI - and finally, its latent disruptive degree.

Keywords: Artificial intelligence, Information Technology, Ethics in AI, Regulation of AI, Generative AI, ChatGPT, Work.

SUMÁRIO

INTRODUÇÃO	11
CAPÍTULO 1: REFERENCIAL TEÓRICO E METODOLOGIA	14
1.1 Inovações e difusão da Inteligência Artificial.....	14
1.2 Governança na Inteligência Artificial.....	17
1.3 Mercado de trabalho.....	18
1.4 Metodologia.....	20
CAPÍTULO 2: DESENVOLVIMENTO DA INTELIGÊNCIA ARTIFICIAL	22
2.1. Processo evolutivo da Inteligência Artificial.....	22
2.2. Difusão atual	24
2.3 Empresas destaques em aplicação de Inteligência Artificial.....	27
2.4 Riscos em relação à adoção de Inteligência Artificial.....	29
2.5 Expectativas do uso no futuro.....	30
CAPÍTULO 3: DESAFIOS ÉTICOS E MORAIS DA INTELIGÊNCIA ARTIFICIAL	41
3.1 Relevância da atuação governamental em benefício da sociedade.....	41
3.2 Incertezas sobre a Inteligência Artificial.....	43
3.3 Danos éticos e legais gerados pela Inteligência Artificial.....	46
3.4 Casos de uso ético da Inteligência Artificial em benefício da sociedade.....	49
3.5 Posição dos Governos em relação a Inteligência Artificial.....	51
3.5.1 Estados Unidos.....	52
3.5.2 Europa.....	56
3.5.3 China.....	58
3.5.4 América Latina.....	59
CAPÍTULO 4: INTELIGÊNCIA ARTIFICIAL GENERATIVA	63
4.1 Funcionamento da Inteligência Artificial generativa.....	63
4.2 Difusão atual.....	65
4.3 Riscos e desafios éticos.....	70
CAPÍTULO 5: IMPACTOS DA INTELIGÊNCIA ARTIFICIAL NO MERCADO DE TRABALHO	76
5.1 Nível de trabalho e influência da Inteligência Artificial.....	77

5.2 Diferença de impactos da Inteligência Artificial relativos ao nível de desenvolvimento do país e a qualificação da mão-de-obra.....	81
5.3 Possíveis cenários resultados do fluxo disruptivo da Inteligência Artificial.....	84
5.4 Potencial da Inteligência Artificial generativa.....	87
5.5 Setores de tecnologia impulsionados pela Inteligência Artificial generativa.....	89
5.6 Impacto do nível de trabalho pela Inteligência Artificial generativa.....	90
5.7 Variação do nível de emprego atual.....	93
CONSIDERAÇÕES FINAIS.....	99
REFERÊNCIAS BIBLIOGRÁFICAS.....	103

LISTA DE GRÁFICOS

Gráfico 1 – Pedidos de patentes de IA por área geográfica, 2010 – 2021 (% total do mundo).....	34
Gráfico 2 – Somatório de investimentos privados em IA de 2012 – 2022 (US\$ milhões).....	35
Gráfico 3 – Investimento privados em IA de 2012 – 2022 (US\$ milhões).....	36
Gráfico 4 – Número de publicações relacionadas a IA por país de 2000 – 2022.....	37
Gráfico 5 – Publicações relacionadas a IA por setor, 2010 – 2021 (% total do mundo)	38
Gráfico 6 – Produção global de semicondutores por país, 1990 – 2030 (% total do mundo)	40
Gráfico 7 – Número de incidentes e controvérsias envolvendo IA de 2012 – 2021.....	45
Gráfico 8 – Número de casos legais relacionados a IA nos EUA de 2000 – 2022.....	55
Gráfico 9 – Investimentos privados em <i>startups</i> de IA generativa, 2021 – 2023 (US\$ bilhões).....	66
Gráfico 10 – Evolução do número de usuários nas principais empresas de tecnologia.....	67
Gráfico 11 – Nível percentual de desemprego por região de 1991 – 2023.....	78
Gráfico 12 – Nível percentual de emprego por região de 1991 – 2023.....	79
Gráfico 13 – Intensidades de mão-de-obra em manufatura por qualificação de 1996 – 2009.....	82
Gráfico 14 – Nível percentual da distribuição da carga de trabalho ocupacional exposta à automação por IA, 2023.....	91
Gráfico 15 – Nível percentual de exposição a tecnologia GPT por grupo de emprego.....	92
Gráfico 16 – Número de demissões na área de TI por trimestre, 2020 – 2023.....	94
Gráfico 17 – Índice de postagens de emprego no Indeed de fev. 2020 – abr. 2023.....	96
Gráfico 18 – Empregos nas grandes empresas de tecnologia antes e depois das ondas de cortes de gastos, 2018 – 2023.....	96

LISTA DE TABELAS

Tabela 1 – ChatGPT em comparação com os principais mecanismos de busca em dezembro de 2022, em milhões.....	68
Tabela 2 – Estrutura de custo do mecanismo de busca do Google e simulação da performance do ChatGPT caso substituísse esse mecanismo, em dólares, 2023.....	68

INTRODUÇÃO

Em meados do século XX inicia-se a história da inteligência artificial (IA). É nesse período que os cientistas começam a discutir fortemente a ideia de que seria possível construir um mecanismo digital, com capacidade de executar funções antes realizadas apenas pela inteligência humana. A IA pode ser definida como um conjunto de ferramentas com o objetivo de desempenhar funções – em diferentes níveis – anteriormente feita apenas por humanos. Estas são compostas por um agregado de dados, potência de processamento e algoritmos sofisticados, podendo manipular rapidamente um volume muito maior de informações do que um humano. Além de que, pela sua capacidade de aprendizado autônomo, quanto mais ela atua, mais conhecimento ela acumula (ANYOHA, 2017; MIT TECHNOLOGY REVIEW, 2023).

O foco principal da IA no contexto atual é, por meio de seus *softwares* altamente capacitados, fornecer inteligência para máquinas, otimizando processos por meio de robótica, geração de dados e automação. Além disso, as IAs mais recentes e desenvolvidas também podem buscar assumir posições mais complexas, como formação de opinião, entendimento de senso comum e comportamento social (GANTORI, 2017).

Um dos principais pretextos para essa linha de pensamento foi a criação do primeiro computador eletrônico digital em 1946, o ENIAC, que originou o desenvolvimento de computadores. Posteriormente e mais importante ainda, em 1950, o matemático Alan Mathison Turing realizou um estudo questionando a capacidade de pensamento das máquinas. Essa tese foi chamada de “jogo da imitação”, em que era testado se uma máquina poderia replicar respostas semelhantes às de um humano perante as mesmas perguntas, sendo esse estudo o ponto inicial de maior relevância para discussão de IA. (TURING, 1950).

O processo de desenvolvimento da computação, que está atrelado com o progresso da IA, é composto por vários testes e pesquisas publicados nos anos seguintes à teoria proposta por Alan Turing. Assim, sem delongar em detalhes históricos, nas décadas posteriores a esse estudo o nível de processamento dos computadores naturalmente foi progredindo, e em conjunto a sua capacidade de executar funções e resolver problemas.

Nas décadas subsequentes esse desenvolvimento começou a se manifestar, tiveram eventos que acentuaram esse desenvolvimento. Um deles foi a elaboração da lei de Moore em 1965, que determinava que o número de transistores dos chips iria dobrar a cada dois anos, mantendo o mesmo custo, representando um marco para a computação (ANYOHA, 2017).

Já em 1997, um grande marco para a IA foi o Deep Blue, um supercomputador equipado com um *software* desenvolvido pela IBM, programado para jogar xadrez e capaz de analisar aproximadamente 200 milhões de posições por segundo. Com isso o então campeão mundial de xadrez, Garry Kasparov, considerado o melhor jogador de todos os tempos, jogou seis partidas contra o algoritmo, ganhou três, empatou duas e perdeu uma. Esse acontecimento passou a ser o primeiro jogo de xadrez em que um computador venceu um campeão do mundo, evidenciando que uma IA de alta performance conseguiu competir com o raciocínio de um humano, bem como retratado no *site* da IBM (2011). No entanto, na época surgiram várias polêmicas sobre o assunto, questionando a validade dessa partida, justamente pela falta de informação do público geral e as incertezas envolvidas no contexto da IA (SILVER, 2015).

No século XXI, o desenvolvimento das máquinas foi muito maior que nos períodos anteriores, principalmente devido à capacidade de processamento que cresce exponencialmente nos termos da lei de Moore. O número de transistores dentro de um microchip, que é uma das medidas de performance dos computadores, de 1970 a 2000, passou de mil para 50 milhões, já nesse século, em 2020, já passou a marca de 50 bilhões. Como resultado, grande parte dos limites de processamento que os cientistas enfrentavam no século passado foram quebrados com o progresso da computação, acelerando ainda mais a capacidade da IA (ANYOHA, 2017; KELLEHER, 2022).

Entretanto, a parte ética e os efeitos da implementação da IA na sociedade não tiveram o mesmo nível de desenvolvimento, principalmente inerente às políticas relacionadas à IA. Sem regulamentação ou transparência, algoritmos de automação são usados para tomar decisões importantes que alteram vidas, sem uma supervisão apropriada (FONSEKA, 2017).

Nos dias atuais, devido a sua capacidade de realizar atividades humanas, a IA está difundida em quase todos campos da sociedade. Neste sentido, a IA afeta vários setores diretamente, se comportando de maneira diferente de muitas outras inovações tecnológicas, que atingem apenas áreas específicas. A tendência é que a IA, nos próximos anos, se fará presente na maioria dos ramos de conhecimento, com uma participação sólida nas esferas das finanças, governo, saúde, indústria e comunicação (SCUDILIO, 2020).

Um pretexto para a discussão sobre a difusão da IA não ser maior, é que ela muitas vezes não é perceptível para seus usuários, mesmo esse tipo de tecnologia sendo usada diariamente nos *smartphones* e computadores. O conceito de IA tende a evocar imagens de computadores e robôs autoconscientes, como nos filmes “Eu, Robô”, “Wall-E” e “O Exterminador do Futuro”, mas por enquanto essa tecnologia chamativa, na maioria das vezes

está restrita à ficção e ao cinema. Já em outros casos, a IA não tem tanta visibilidade por falta de transparência das empresas que a utilizam, agregado à ausência de regulação, gerando impactos no meio socioeconômico, sem serem percebidas (FONSEKA, 2017).

Assim, o potencial disruptivo da IA não é muito claro para o senso comum, mas é altamente significativo. A tecnologia já implicou em mudanças imutáveis, e muitas ainda vão se manifestar. O elevado desenvolvimento da tecnologia em conjunto com suas consequências positivas e negativas, leva ao questionamento desta monografia, que é de que maneira a IA influencia a sociedade e economia. Na tentativa de acompanhar o desenvolvimento da tecnologia, muitos cientistas e estudiosos fazem pesquisas e levantamentos sobre os impactos da IA, com dados e questionamentos importantes sobre a tecnologia, os quais foram usados na elaboração deste trabalho.

O objetivo geral do estudo é entender os principais pontos que envolvem a IA, e a partir deles mostrar ou prever como o seu potencial transformador opera na economia, e de que forma ele se destaca de outras inovações tecnológicas. A hipótese a ser defendida associa a evolução da IA – desenvolvendo setores e gerando novas profissões – a seus impactos disruptivos na estrutura de mercado, e isso, por sua vez, desperta uma necessidade de mudanças de estratégia pelas companhias e de regulação pelo governo, distinto do cenário vigente. A fim de testá-la, serão analisadas variáveis como difusão da tecnologia, questões éticas e normativas, qualificação dos trabalhadores e variação do nível de emprego dos setores.

A estruturação desta monografia conta com cinco capítulos: o primeiro capítulo mostra o contexto das teorias evolucionistas em consonância com a IA, evidenciando a contribuição dos principais autores usados neste estudo. O segundo capítulo discorre sobre o processo de desenvolvimento da tecnologia, destacando sua difusão, riscos de adoção e expectativas de uso futuro. O terceiro analisa os desafios éticos e morais na ampliação e implementação da IA, e apresenta alguns marcos legais da tecnologia vinculados às incertezas presentes na mesma. Já o quarto capítulo apresenta o funcionamento da IA generativa – atual modelo de maior impacto – e as peculiaridades que a diferem de outras IAs. Por fim, o quinto capítulo apresenta os impactos da IA no mercado de trabalho, a variação de nível de emprego na própria área da tecnologia e em outros setores e, enfim, as considerações finais sintetizam as principais conclusões deste estudo.

CAPÍTULO 1: REFERENCIAL TEÓRICO E METODOLOGIA

1.1 Inovações e difusão da Inteligência Artificial

Segundo Schumpeter (1961), o impulso fundamental para o progresso econômico em uma economia capitalista não é apenas a acumulação de capital ou a eficiência na produção, mas sim a inovação e a introdução de novas ideias e tecnologias no mercado. Essas inovações têm o potencial de interromper e substituir antigas formas de produção e organização. Assim, este processo de destruição criadora é essencial para o desenvolvimento econômico a longo prazo, e é dele que se constitui o capitalismo e a ele devem-se adaptar as empresas.

Em relação à teoria inovativa, os autores neo-schumpeterianos, também denominados como evolucionários, fundamentam-se na ideia defendida por Schumpeter de que mudanças tecnológicas impulsionam o desenvolvimento capitalista. De acordo com essa teoria, as inovações são formadas por meio de uma interação complexa entre alguns fatores econômicos fundamentais (busca por novas oportunidades de lucro e novos mercados, tendência à redução de custos e automação, etc.), juntamente com fatores institucionais (interesses e estrutura das empresas existentes, efeitos das agências governamentais, etc.) (VIEIRA, 2010; DOSI, 1982).

Segundo Pavitt (1984), nesse meio teórico, é reconhecido por uma ampla gama de estudiosos que a produção, adoção e disseminação de invenções tecnológicas são fatores essenciais no desenvolvimento econômico e na mudança social. Dessa forma, a inovação técnica é uma característica distintiva dos produtos e setores industriais dos mercados mais competitivos. Para Dosi (1982), o surgimento desse processo inovativo é atribuído às instituições que produzem e direcionam a acumulação de conhecimento e experiência, e pela existência de uma multiplicidade de atores dispostos a assumir riscos, prontos para experimentar soluções técnicas e comerciais diferentes.

No entanto, essas representações dos processos de mudança de tecnologia encontradas na economia se contrapõem à abordagem neoclássica. Assim, isso ocorre à medida que sustenta a inovação como papel fundamental no processo dinâmico da economia, opondo-se ao conceito deles de análise estática e de equilíbrio otimizado da firma. Na formulação neoclássica original, a nova tecnologia se difunde instantaneamente por todo o capital. Já nessa nova onda de pensadores neo-schumpeterianos, a tecnologia está associada ao capital que a incorpora e, portanto, a adoção de uma nova técnica é impulsionada e limitada pela sua taxa de investimento (VIEIRA, 2010; PAVITT, 1984).

Dessa forma, além de ser uma teoria com enfoque na inovação, esse lado mais evolucionário é parcialmente resultado de uma crescente conscientização de que: a teoria neoclássica não consegue lidar adequadamente com a dinâmica de desequilíbrio envolvida em certos tipos de competição de alguns setores. De forma mais ampla, falha em lidar com os processos de crescimento econômico impulsionados pela mudança tecnológica, principalmente em áreas intensivas em pesquisa e desenvolvimento (P&D), como na área de farmacêuticos ou computadores (NELSON & WINTER, 2002).

Pavitt (1984) também compartilha desse pensamento, que nesses setores baseados em ciência as principais fontes de novas tecnologias são as atividades de P&D, com base no rápido desenvolvimento nas áreas científicas subjacentes nas universidades e em outros campos. A ampla gama de aplicações com base nessas tecnologias tem levado a um rápido crescimento das firmas bem-sucedidas e inovadoras nesses setores, e elas têm tido pouco incentivo para buscar oportunidades inovadoras além de seu setor principal. Além disso, dada a sofisticação das tecnologias e das bases de conhecimento adjacentes, tem sido difícil para empresas de fora desses setores ingressarem neles.

Nesse sentido, os vínculos que ligam a inovação ao crescimento e tamanho das firmas são centrais para a pesquisa sobre a dinâmica da competição schumpeteriana por Nelson e Winter (1982). Eles fazem uma previsão sobre esses setores, com taxas rápidas de mudança tecnológica, incerteza nos resultados de investimentos em inovação e com fortes possibilidades das empresas se apropriarem da vantagem inovadora. O resultado são altas tendências de, ao longo do tempo, surgirem concentrações tanto da produção quanto das atividades inovadoras (PAVITT, 1984).

Assim, a partir desse processo descrito pelos pensadores evolucionistas são desenvolvidas várias inovações, incluindo a IA. No entanto, existem diferenças entre mudanças tecnológicas anteriores em relação à IA. Por mais que muitas características inovativas se assemelham às outras, a IA possui aspectos muito peculiares vindo de um processo inovativo que cresce exponencialmente. Diferente de outros processos de automação estudados por autores passados, a característica de auto aprendizado e possibilidade da geração de conteúdo próprio concedem uma certa autonomia à tecnologia, fugindo da realidade de épocas passadas. Com isso surge uma necessidade intrínseca de analisar estudiosos contemporâneos e especialistas na área.

A mudança de trajetória acontece tão rápido que previsões feitas em estudos na década passada, por mais que respeitadas e de autores renomados em IA, também já fogem da realidade

atual. Um exemplo muito claro disso foi nos estudos de profissões mais suscetíveis a serem impactados pela IA. Na década passada contavam em sua maioria com profissões mecânicas, rotineiras e de baixa qualificação. Porém, com a difusão da IA generativa no final de 2022, capaz de criar novos conteúdos próprios, profissionais da área criativa e com melhores qualificações e salários, se tornaram alvos mais vulneráveis à tecnologia como mostra o estudo mais recente de Eloundou *et al* (2023) apresentado no quinto capítulo.

Nesse sentido, levando em conta o processo inicial de difusão, Grennan *et al* (2022) discutem como a adesão de IA nas empresas deve acontecer junto a um entendimento da tecnologia, tanto para os desenvolvedores quanto pelos usuários do produto. As pessoas usam aquilo que entendem e confiam. Logo, as companhias que facilitam o uso da tecnologia terão vantagens competitivas, não apenas com os usuários de suas organizações, mas também com reguladores e consumidores. No entanto, a complexidade e dificuldade que envolvem esses processos da IA leva a uma falta de compreensão da mesma.

Chui (2022) não só faz uma análise sobre o processo inicial da tecnologia de IA generativa, mostrando a importância da adaptação ao novo contexto pelas empresas, como faz um estudo comparativo entre as que são destaques na adoção antecipada de IA em relação às outras. O autor explica de que forma essas companhias conseguem se diferenciar e mostra as vantagens competitivas construídas por elas, dentro das especificidades da IA. No âmbito geral, Nelson e Winter (2002) explicam esse processo, em que as grandes firmas são propensas a gastar mais em P&D do que as menores. Logo, a inovação bem-sucedida tende a aumentar a lucratividade de uma empresa, em uma quantia absoluta proporcional à sua escala, e, portanto, a fazer com que a organização crescesse e gastasse mais em P&D.

Em relação às partes mais técnicas da IA, Patel & Ahmad (2023) fazem um estudo de caso referente à aplicação da IA generativa no contexto atual, em que seu nível de eficiência inicial ainda é baixo. Nesse sentido, a difusão em massa dessa tecnologia nas condições vigentes, o produto final seria mais desejável que o dos concorrentes, mas ineficaz em relação ao custo fixo adicional gerado. Pela necessidade de ampliar completamente o sistema computacional para sustentar esse tipo de algoritmo, o valor de investimento e os gastos em *hardware* seriam insustentáveis no momento.

Quanto aos riscos dessa tecnologia, que tem um potencial de interferência não só diferente, mas maior do que das IAs anteriores, Korda (2023) faz uma análise de como esses algoritmos podem trazer riscos para a sociedade. Em seu estudo, o autor mostra como ele, e outras pessoas pelo mundo inteiro, conseguiram achar brechas na tecnologia para fazerem um

uso não ético da mesma. Assim, situações de oportunismo podem surgir nesse contexto, e há uma necessidade de um maior comprometimento das empresas antes de implementarem suas IAs, e do governo em regular. Existem riscos e o poder público precisa ser ativo nesse meio.

1.2 Governança na Inteligência Artificial

Segundo Nelson e Winter (1982), governos e políticas são uma parte importante do ambiente que molda a evolução das atividades do setor privado:

- Leis e normas relacionadas ao que é patenteável, e sobre acordos de licenciamento aceitáveis ou obrigatórios, influenciam as vantagens relativas de inovar e imitar;
- A legislação antitruste e sua interpretação administrativa e judicial definem comportamentos competitivos aceitáveis;
- Regimes regulatórios restringem e exigem certas ações privadas;
- Sistemas públicos de ensino e programas de apoio educacional influenciam o fluxo de pessoal capacitado para atividades de pesquisa e desenvolvimento;
- Programas de apoio governamental à P&D têm fornecido, desde a Segunda Guerra Mundial, aproximadamente metade do financiamento total para pesquisa e desenvolvimento;

Uma parte significativa da atividade econômica é conduzida por organizações públicas, em vez de privadas. A evolução das capacidades econômicas e comportamentais deve ser compreendida como uma economia mista. Assim, as políticas públicas evoluem parte em resposta às mudanças nas demandas e oportunidades, podendo resultar de inovações tecnológicas ou variações nas estruturas de mercado. Além disso, as políticas públicas podem refletir não apenas mudanças nas condições objetivas, mas também trocas nos valores ou na compreensão (NELSON & WINTER, 1982).

Embora muitas discussões políticas sejam conduzidas em termos de busca pelo interesse público, esse "público" é mais uma figura de linguagem do que uma entidade concreta com metas explícitas. O "público" real, interessado nas escolhas políticas e nos resultados, possui uma composição diversa e divergente, com interesses que estão, pelo menos em parte, em conflito (NELSON & WINTER, 1982).

Além disso, existem várias maneiras pelas quais as partes interessadas podem influenciar a formulação de políticas. A política e a administração da defesa claramente divergem da educação, que por sua vez difere da tecnologia. Muitas questões de política pública são complexas, a natureza dos problemas e das opções não são bem compreendidas, e os valores em jogo estão longe de ser transparentes (NELSON & WINTER, 1982).

No contexto da IA, Larsen (2022) faz uma cobertura sobre o processo atual de disputa entre os Estados Unidos e União Europeia com a China em relação ao domínio do mercado de IA. O autor aponta a posição dos Governos em relação à tecnologia, mostrando tanto os motivos que impulsionam esse conflito, como as críticas e justificativas usadas pelos países para as medidas protecionistas empregadas. De fato, os EUA e UE apontam preocupações em relação à China usar IA em um sistema de pontuação social com o objetivo de implementar controle em larga escala. A justificativa, que não necessariamente representa o real motivo, é que esse programa impõe uma ameaça às liberdades fundamentais, suprimindo a liberdade de expressão pela vigilância arbitrária ou ilegal.

Larsen (2022) também mostra o processo inicial das medidas reguladoras de IA adotadas por esses países líderes em relação à tecnologia. Além dele, Engler (2021, 2022, 2023) aborda os motivos e ideais desses principais países em fazer suas escolhas referentes à regulamentação, por meio de suas políticas, planos e leis. Esse processo acontece desde de forma protecionista com o controle majoritário do governo na China, até liberal como nos EUA, cada um contendo suas particularidades. O autor também examina questões relacionadas à ética e riscos relacionados à IA, e mostra casos de usos em que a tecnologia é prejudicial. Assim, ele evidencia a importância do governo em interferir nesse processo para garantir que a tecnologia não cause danos à sociedade.

Nesse mesmo contexto, Korinek e Balwit (2022) mostram a relevância do governo nesse processo, explicando o conceito de alinhamento social, em que os objetivos da IA devem condizer com as normas da sociedade, para que não gerem externalidades e prejudiquem as pessoas afetadas por ela. Assim, a dificuldade se encontra não só na complexidade da tecnologia como apontado por Grennan (2022), mas pelo efeito de escala, sendo que, quanto maior o grupo de pessoas, como em países ou continentes, mais difícil é alinhar a tecnologia para beneficiar a todos.

1.3 Mercado de trabalho

Na visão de Dosi (1982), um critério importante relacionado a inovações é a sua capacidade de redução de custos e, por consequência, seu potencial de poupar trabalho. Segundo ele, isso está em consonância com a sugestão de Nelson e Winter sobre as trajetórias naturais em direção à mecanização e exploração das economias de escala.

Além disso, em sociedades onde existe um conflito entre questões industriais e sociais, a substituição de máquinas por mão-de-obra deve ser um poderoso determinante de qual será o

foco no processo de busca por novas tecnologias. De maneira mais geral, esse confronto entre industrial e social têm a tenência de definir o processo de seleção de novos paradigmas tecnológicos. Isso implica tanto critérios negativos, em possíveis exclusões de mão-de-obra, quanto como critérios positivos, na demanda por novas profissões. Nesse sentido, pode-se definir uma relação de longo prazo entre padrões de desenvolvimento social e as áreas de inovações escolhidas (DOSI, 1982).

No âmbito da IA, Littman *et al* (2021) elaboram um estudo complexo dos impactos da IA por várias óticas, sendo um deles as implicações da mudança tecnológica no mercado de trabalho. A questão levantada é que perante a história inovações trouxeram tanto benefícios quanto malefícios à sociedade. O potencial da IA para replicar o trabalho humano a um custo mais baixo levanta preocupações sobre seu impacto no bem-estar dos trabalhadores, mas é difícil de caracterizar precisamente qual será o nível disso. Para ele, até o momento, a IA ainda não foi responsável por grandes efeitos econômicos agregados.

De acordo com Baily e Korinek (2023), a IA não apenas torna os profissionais mais produtivos, como aumenta a taxa de inovação, estabelecendo as bases para uma aceleração significativa no crescimento econômico. Como uma tecnologia de propósito geral, a IA terá impacto em uma ampla gama de setores, estimulando investimentos em novas habilidades, transformando processos empresariais e alterando a natureza do trabalho. Os avanços rápidos podem trazer grandes benefícios, mas também podem gerar riscos significativos, portanto, se faz necessário que o progresso seja melhor orientado para evitar externalidades à sociedade.

Finquelievich (2019) aponta que o progresso da substituição de empregos por IA não segue uma trajetória linear, existem vários possíveis cenários futuros. Mesmo que ocorram substituições, isso aumentará a composição orgânica do capital e a produtividade, mas não necessariamente eliminarão o trabalho humano, uma vez que é a fonte de ganho de capital necessária para a subsistência da sociedade. Segundo ela, os avanços alcançados se reduzem a três elementos fundamentais: progresso nos algoritmos, aumento do poder computacional, e maior acesso aos dados. A combinação desses fatores determinou o rápido avanço da IA nos últimos anos.

Em relação aos países em desenvolvimento, Korinek e Stiglitz (2021) afirmam que o progresso na IA ameaça reverter os ganhos que mercados emergentes tiveram ao se integrar à economia mundial nas últimas décadas, e por consequência agravar a pobreza e a desigualdade. As inovações têm a tendência de gerar desemprego, principalmente em países onde a mão-de-obra é menos qualificada. Logo, são propensos a favorecer os países desenvolvidos criadores

dessas tecnologias. Assim, são necessárias políticas que mitigariam esses efeitos adversos sobre as economias emergentes, e ao mesmo tempo compartilhar os benefícios da IA de forma mais ampla com esses países.

Autor *et al* (2022) mostram que existem muitos sinais de que a indústria manufatureira está perdendo força. Historicamente, países conseguiram realocar sua força de trabalho da agricultura para a indústria, e gerar ganhos significativos de produtividade em toda a economia. No entanto, nos poucos países de baixa renda onde a industrialização parece não ter perdido força, sua qualidade é muito baixa. Os autores atribuem essa mudança principalmente aos impactos das inovações tecnológicas na mão-de-obra, como a automação. Além disso, as evidências indicam que o efeito de perda foi mais forte para os trabalhadores menos qualificados, mais presentes em economias emergentes, que se desejam ser competitivos, precisam importar as inovações dos países desenvolvidos.

Já Ewunetie e Hofheinz (2022) investigam a relação entre qualificação dos profissionais e inovações, em que os empregos que requerem altos níveis de treinamento, experiência e educação aumentaram 68% entre 1980 e 2015. Os empregadores relatam um nível sem precedentes de escassez de habilidades e dificuldades na contratação, mesmo com o número de pessoas com diplomas universitários estar crescendo rapidamente. Para eles, a inovação tecnológica frequentemente traz consigo desequilíbrios de habilidades, pois exige pessoas com conhecimento e paciência para usar as novas ferramentas. Tecnologias complexas estão chegando ao mercado de trabalho muito mais rápido do que antes, e nesse dinamismo, o estudo contínuo e o reaprendizado constante são a norma, e não a exceção.

Korinek and Juelfs (2022) apresentam diferentes resultados da aplicação da IA no mercado de trabalho. Os autores discutem o papel das instituições para tornarem a sociedade resistente a um futuro em que máquinas autônomas, cada vez mais inteligentes, substituam o emprego humano e reduzem os salários. Eles apresentam preocupações decorrentes desse cenário, em que a IA implica a redundância econômica do trabalho, logo seria necessário realocar o emprego e renda, caso isso se materializasse.

1.4 Metodologia

O método de procedimento escolhido para a condução da monografia foi uma intensa revisão bibliográfica sobre o tema, que procedeu da pesquisa de natureza, majoritariamente, qualitativa. O intuito disso é permitir uma ampla avaliação do cenário da IA, que além de complexo é muito vasto, por poder ser aplicada na maioria dos setores econômicos e sociais.

Para isso, é necessária uma investigação de várias matrizes e possibilidades, na tentativa de entender e qualificar os padrões encontrados na tecnologia e no comportamento dos usuários.

Por meio de uma revisão sistemática da literatura, o propósito é fornecer uma visão geral do conhecimento, feita a partir de fontes secundárias resumidas e as articulando de uma maneira acessível, visto que pela complexidade do tema muitas vezes ele é desencorajador de ser estudado. Dado que a finalidade da pesquisa é exploratória, visando a proporcionar maior familiaridade com o assunto a partir do esclarecimento de conceitos e busca por padrões, o estudo procura indicar os problemas e hipóteses sobre o tema. O método de abordagem utilizado é o método indutivo, partindo de análises parciais, tentando alcançar uma conclusão geral.

A monografia olha de maneira abrangente as características e competências da IA, uma vez que tem como objetivo discorrer sobre os possíveis impactos da tecnologia na economia, que por ser um assunto relativamente recente requer mais investigações para se obter um conhecimento minucioso. Abordar somente sobre a difusão da IA e a sua implicação no mercado de trabalho não seria prudente, visto que os riscos e incertezas que envolvem a sua questão ética estão intrinsecamente ligados a seus impactos socioeconômicos. Nesse mesmo sentido, não se atentar para as regulações que a envolvem e sua característica mais promissora (IA generativa) deixariam o trabalho inconsistente, uma vez que a análise é sobre suas possibilidades – por se tratar de uma inovação com muito potencial, efeitos ainda pouco previsíveis e alta intensidade de mudança – e esses fatores influenciam diretamente no seu futuro.

CAPÍTULO 2: DESENVOLVIMENTO DA INTELIGÊNCIA ARTIFICIAL

2.1. Processo evolutivo da Inteligência Artificial

Os impactos da IA variam de acordo com seu tipo e área de implementação, podendo tanto gerar a substituição de empregos de outras áreas, quanto apenas complementá-los. Essas características são muito marcantes da IA, justamente por otimizar um processo e aumentar a produtividade em uma parte da cadeia produtiva, basicamente resultando em mais benefícios com um menor esforço. Um exemplo é na agricultura, em que semeadoras e colheitadeiras são automatizadas com *Global Positioning System* (GPS) e piloto automático, e assim distribuem fertilizantes em enormes plantações ajudando a manter os custos operacionais baixos. Esse tipo de IA substitui o trabalho de pessoas, e quanto maior a plantação, maior o desemprego gerado. Mas como a aceitação da tecnologia é alta devido a sua viabilidade, a tendência desse processo é só aumentar. Os gastos globais em tecnologias e sistemas agrícolas inteligentes devem triplicar em receita até 2025, atingindo US\$ 15,3 bilhões. Já os gastos específicos em tecnologias e soluções de IA tendem a crescer de US\$ 1 bilhão em 2020 para US\$ 4 bilhões em 2026, atingindo um *Compound Annual Growth Rate* (CAGR), taxa de crescimento anual composta, de 25,5% (COLUMBUS, 2021).

Por outro lado, a IA em um algoritmo de busca, como na Netflix, que recomenda títulos de filmes e séries baseado nas pesquisas do usuário, não substitui o trabalho de outra pessoa, pois o mesmo não seria feito manualmente para a conta de cada usuário. Nesse caso a IA cria um novo processo que gera benefícios para o produto final e emprego para a pessoa que gerencia o algoritmo. Logo, diferentes tipos de IA promovem impactos distintos.

Outra situação é quando um processo deveria ser desempenhado manualmente, mas por diversas razões, como um alto custo, a atividade deixa de ser realizada, é feita de forma ineficiente ou é de forma incompleta. Um exemplo desse tipo de atividade, é a fiscalização de fraudes fiscais relacionadas à taxação sob residências na França, que resultou em um algoritmo experimental de IA como solução. Ele foi implantado em regiões que a tributação de casas representa uma grande parte da arrecadação – como Ville Rennes, onde 70% dos impostos providos vêm de residências – junto com outros oito departamentos franceses de um total de 101. Assim, a AI analisa imagens obtidas por drones para detectar piscinas nas residências e identifica as que não são devidamente reconhecidas e tributadas, no qual somente nesses oito departamentos mais de 20.000 piscinas não foram declaradas, representando 5,7% das atualmente legalizadas. Em valores, a tributação dessas piscinas resultaria em cerca de €10

milhões em receitas fiscais adicionais para 2022, a estimativa é que esse valor alcance €40 milhões em 2023. Isto mostra novamente que diferentes tipos de IA promovem impactos distintos (MOODYS, 2022).

Um mesmo processo de uma IA pode gerar diferentes impactos em áreas divergentes, por exemplo um aplicativo de reconhecimento facial. Ao ser usado no telefone para desbloquear a tela, a inovação somente substitui o processo anterior que era digitar a senha pelo teclado, assim mudando o papel de uma tecnologia. Por outro lado, quando uma empresa que tem um porteiro para regular a entrada no estabelecimento, e implementa um sistema automático de reconhecimento facial que libera a entrada para as pessoas cadastradas, a IA substituiu a função de humanos. Dessa forma, uma tecnologia com a mesma função de reconhecer a face de um ser humano, ao ser implementada de forma distinta em diferentes áreas, gerou consequências divergentes. Ainda, a tecnologia em si pode esbarrar em outros problemas, essa mesma tecnologia pode ser utilizada para vigilância e monitoramento, reforçando uma falta de liberdade e invasão de privacidade, gerando vários questionamentos éticos que serão discutidos no terceiro capítulo desta monografia.

Além disso, os impactos da IA também variam de acordo com o local e o grupo demográfico que habita nele. Usando um exemplo simples, a IA na agricultura citada acima, ao ser implantada em um país com pouca atividade agrícola não acarreta resultados tão volumosos. Já no Brasil que tem uma grande quantidade de terras produtivas o impacto gerado por essa inovação tecnológica seria maior. Mas a realidade é que o local e o grupo demográfico brasileiro ainda não são propícios para isso, devido ao baixo investimento em IA no país em relação a outros. Assim, nações europeias, mesmo com um território produtivo menor, possuem uma exposição muito maior que o Brasil em tecnologias de IA na agricultura. O crescimento de IA no campo citado anteriormente, refletirá muito mais nos países desenvolvidos e intensivos em tecnologia, tendo o EUA como maior utilizador atual, e a UE com a maior previsão de crescimento (MORDOR INTELLIGENCE, 2021).

Um exemplo mais específico é na China que, por ser uma das líderes em adoção de IA, tem a possibilidade de digitalizar os supermercados de ponta a ponta, desde o atendimento até o estoque, podendo aumentar a eficiência e as margens de lucro, reduzir os custos, agradar os clientes e obter ganhos em relação à concorrência. No entanto, esse processo está atrasado em relação aos outros países, mesmo com o mercado varejista de alimentos chinês crescendo cerca de 3% ao ano entre 2015 e 2020. Os 20 maiores operadores de alimentos na China respondem por apenas 15% do mercado, já no Ocidente, os cinco maiores representam entre 40% e 60%.

Essa fragmentação gera menor eficiência e incapacidade de agir rápido perante as mudanças de tendências e decisões dos consumidores, que explica o não sucesso ou atraso dos chineses nas automações nessa área em específico (JACOBSEN *et al.*, 2022).

Dessa forma, a localização geográfica também influencia na extensão do impacto gerado pela IA, e essas diferenças em termos de implicações, juntamente com outros critérios relacionados à IA, serão o objeto de análise neste capítulo. Com base em diferentes modelos de IA, juntamente com suas características como localidade, ocupação e grupo demográfico, tendo atenção nos critérios mais notórios, e os padrões, caso existam, que podem ser percebidos entre eles. E, por fim, entender as previsões de avanços em IA e tecnologias relacionadas, que podem, a partir de seus diversos impactos, representar um ponto de virada crítico na história da economia.

2.2. Difusão atual

No contexto atual, em que a Tecnologia da Informação (TI) já passou por vários tipos de processos de difusão, outras tecnologias inovativas derivadas desse processo de digitalização passam a se destacar, como armazenamento em nuvem, Internet das Coisas e principalmente a IA. Assim, a indústria 4.0 – também chamada de quarta revolução industrial – é a próxima fase no progresso dessas tecnologias, impulsionada por tendências disruptivas, incluindo o aumento de dados e análises, conectividade, interação do humano com as máquinas e melhorias na robótica. Como consequência, algumas dessas tecnologias que são disruptivas transformam os requisitos de trabalho, exigindo profissionais com níveis mais elevados de habilidade (MCKINSEY & COMPANY, 2022).

Cada indústria será transformada durante a quarta revolução industrial, mas algumas mais do que outras. A natureza da transição para a Indústria 4.0 diferirá pelos tipos específicos de tecnologia que estão sendo adotados, bem como pela infraestrutura e habilidades existentes das organizações. Uma análise do Instituto McKinsey & Company (2022) mostra que os pioneiros da Indústria 4.0 – instalações bem encaminhadas para adotar a IA e outras tecnologias avançadas até 2025 – podem esperar uma mudança positiva de 122% no fluxo de caixa. Outras empresas usuárias devem alcançar apenas 10%, enquanto as que falharem completamente em adotar a IA podem ver uma queda de 23%.

Durante o processo de implementação dessas tecnologias, muitas companhias enfrentam problemas em identificar os dados mais relevantes ao lidar com algum desafio específico. Depois de décadas coletando informações, as empresas muitas vezes possuem abundância em

dados, mas pouca compreensão sobre eles, tornando quase impossível navegar pelos milhões de registros feitos em busca de informações significativas. Assim, a IA pode acelerar esse processo ao ingerir grandes volumes de dados, e encontrar rapidamente os elementos mais prováveis de serem úteis para resolução de problemas (BORDEN *et al.*, 2022).

As empresas podem ensinar a IA a navegar por documentos carregados de informações, fornecendo ferramentas como dicionários técnicos, para ajudar os algoritmos da IA a entender as relações semânticas entre seus diferentes tipos de dados. O algoritmo pode criar uma rede dinâmica relacionando todos os documentos e dados, e a partir de uma interface, os usuários da companhia podem usar a base de conhecimento para solucionar até os problemas mais específicos. Por fim, o sistema coletaria o *feedback* dos usuários, e assim a IA melhora seu algoritmo e desempenho por conta própria. Com poder computacional suficiente, a IA pode examinar milhões de possibilidades para identificar oportunidades de otimizar os processos das companhias (BORDEN *et al.*, 2022).

Um dos processos que viabiliza a IA é essa sua capacidade de auto aprendizado, possibilitando mais autonomia no seu progresso, que ocorre pelo aprendizado de máquina ou *Machine Learning*. Esse processo permite ao algoritmo a capacidade de aprender sem serem explicitamente programados. Ao analisar os dados disponíveis, por exemplo, dados históricos sobre parâmetros de produção, composição química de matérias-primas, dados de sensores visuais ou câmeras e imagens de satélite, padrões subjacentes são identificados e o sistema se ensina continuamente com base nessas informações (S&P GLOBAL PLATTS, 2017).

Machine Learning alcança resultados expondo o algoritmo a muitos exemplos de soluções corretas, também pode definir um conjunto de regras e permitir que o sistema aprenda por tentativa e erro. Esse acúmulo de dados é essencial para a expansão da IA e tecnologias decorrentes dela, como em áreas de pesquisa, processamento de linguagem, visão computacional e robótica (PERSET *et al.*, 2022).

Esse algoritmo é uma das principais abordagens utilizadas para análise preditiva e prescritiva com IA. O programa consegue aumentar sua eficiência identificando possíveis variáveis e anomalias, sem necessidade de novos investimentos ou mudanças na sua estrutura. Segundo um relatório da S&P GLOBAL PLATTS (2017), seu uso para prever falhas em equipamentos ou otimizar a logística, mostrou capacidade de alcançar ganhos de eficiência de 3%, 5% ou até mais de 10% em quando aplicado em escala.

Por outro lado, *Deep Learning*, ou aprendizagem profunda, é o subconjunto de *Machine Learning* composto por códigos que permitem que o *software* se treine para realizar tarefas,

como reconhecimento de fala e imagem, se espelhando a redes neurais com vastas quantidades de dados. A tecnologia é baseada em rede neural artificial que imita o processo de aprendizado do cérebro humano, sendo esse seu diferencial. *Deep Learning* pode analisar imagens, vídeos e dados não estruturados, de maneiras que *Machine Learning* teria dificuldade ou não poderia fazer (THE ECONOMIST GROUP, 2022; S&P GLOBAL PLATTS, 2017).

Consequente de suas singularidades e necessidades específicas, a implementação da IA ocorre de forma diferente em cada empresa, que em conjunto com a complexidade desses tipos de modelagem apresentados acima, resulta, no geral, em um efeito de atraso entre o surgimento e uso produtivo dessas novas tecnologias. Isso ocorre em paralelo ao aumento da dependência de IA nas firmas, usados para tomar decisões que podem afetar expressivamente os direitos individuais, a segurança humana e as operações comerciais críticas (GRENNAN *et al.*, 2022).

A partir desses fatores, surge o questionamento sobre o nível de compreensão da sociedade em relação à IA. Os sistemas de IA devem ser entendidos tanto pelos que implementam, quanto pelos usuários e pessoas afetadas pelo seu uso. As técnicas de modelagem que atualmente impulsionam muitas aplicações de IA, como aprendizado de máquina e redes neurais, são inerentemente mais difíceis de serem compreendidas pelos seres humanos. Entretanto, já estão sendo aplicadas e usadas na sociedade (GRENNAN *et al.*, 2022).

Por outro lado, existe ainda a necessidade de que essas tecnologias sejam implantadas com maior facilidade, colocar barreiras pode atrasar o processo de evolução da IA e de seus modelos mais avançados. Um exemplo significativo dessa falta de compreensão, e foco apenas na rápida implementação, acontece no meio público. A computação é fundamental para o desenvolvimento de IA e as tecnologias derivadas dela, com isso os políticos vêm aumentando suas estratégias de implantação especializada em IA. Garantir que os países tenham poder computacional suficiente para atender às suas necessidades é fundamental para capturar todo o potencial econômico da IA (PERSET *et al.*, 2022).

No entanto, o *hardware* e a infraestrutura de computação que possibilitam os avanços da IA recebem menos atenção nos círculos políticos do que outros facilitadores, como as bases de dados. Isso enfatiza o questionamento levantado acima sobre a cognição da sociedade em relação à tecnologia. O resultado dessa desinformação é uma grande lacuna, entre aqueles que têm recursos para desenvolver e treinar os modelos de IA em larga escala, acumulando vantagem competitiva, crescimento e ganhos de produtividade, e aqueles que não têm (PERSET *et al.*, 2022).

Nesse sentido, as empresas que obtêm os maiores retornos financeiros com o uso de IA apresentam melhores práticas que permitem a explicabilidade, logo compreensão sobre a tecnologia. A explicabilidade seria a capacidade de expressar por que um sistema de IA tomou uma decisão, fez uma recomendação ou previsão específica. Com isso, corporações que estabelecem uma maior confiança digital entre os consumidores, consequente de práticas que tornam a IA explicável, geram maiores ganhos econômicos (GRENNAN *et al.*, 2022).

Desenvolver essa proficiência requer entendimento de como o modelo de IA opera e os tipos de dados usados para treiná-lo. Mas visto que modelos de IA, como o *Deep Learning*, podem ser muito complexos e sofisticados, o entendimento de como se chegou nos resultados se torna desafiante até mesmo para os desenvolvedores. Os algoritmos de IA ficam mais inteligentes ao longo do tempo, ingerindo dados ininterruptamente, aumentando sua capacidade e, por consequência, sua complexidade (GRENNAN *et al.*, 2022).

Dessa forma, facilitar o entendimento da IA ajuda os profissionais da área de cinco principais maneiras: aumentando a produtividade, construindo confiança e adoção, identificando novas intervenções geradoras de valor, garantindo que a IA forneça valor aos negócios e mitigando riscos. E além disso, mesmo quando não especificamente exigido, as empresas precisarão garantir que qualquer ferramenta de IA usada para tomar decisões esteja em conformidade com as leis antidiscriminatórias aplicáveis, bem como com as leis que proíbem práticas injustas ou enganosas (GRENNAN *et al.*, 2022).

2. 3 Empresas destaques em aplicação de Inteligência Artificial

São considerados líderes na adoção de AI aqueles que investiram em iniciativas de AI e obtêm resultados bem superiores aos demais. Devido aos ganhos obtidos, este grupo reduzido de firmas mostra uma crescente nos seus investimentos em AI, dedicando-se a levá-la à escala. Esse capital é priorizado para aplicações que geram receita, ao invés das que reduzem custos. Mesmo diante de grandes resultados, as mudanças causadas pela AI geram tanto receio quanto expectativa entre os trabalhadores (RANSBOTHAM *et al.*, 2018).

A partir de uma análise da companhia UBS Securities Australia LTD (2020) mostra que, além dos incentivos pelos resultados obtidos, empresas maiores, com margens e reservas de caixa mais altas tendem a investir mais. Como consequência, houve um aumento significativo nas vendas e no emprego por elas, dado que os benefícios da IA acumularam. A consequência disso foi uma dispersão significativa nos ganhos e, portanto, uma oportunidade para gestores dessas gerarem performances superiores. Além disso, as companhias líderes geram mais

benefícios com a IA devido, na maioria dos casos, a acumularem conjuntos muito extensos de dados, que são propensos a gerar algoritmos mais amplos e completos.

Como evidenciado por Pavitt (1984), setores da tecnologia com alto potencial inovativo tendem a gerar não só grandes barreiras de entrada, mas a dominarem o ciclo inovativo da mesma, devido à base de conhecimento já adquirida. Pode-se alegar que exista uma relação direta entre o aumento na concentração da indústria e os investimentos em tecnologias de IA. Empresas com maiores investimentos obtiveram um crescimento mais rápido em receita e redução de custos, e com isso, aumentaram o investimento. Até o momento, esses efeitos positivos estão concentrados entre as líderes, que fazem maiores investimentos e com isso tem melhores resultados, fortalecendo e gerando continuidade nesse processo de hegemonia. Portanto, investir em IA tende a aumentar a escala das organizações mais produtivas e contribui para a ascensão dessas que são destaques (UBS SECURITIES AUSTRALIA LTD, 2020).

Essas firmas têm um desempenho generalizado melhor que as outras, ficando claro que as lacunas entre o alto e baixo desempenho em IA não estão apenas em áreas técnicas. A diferença também acontece nos aspectos humanos, como o alinhamento dos executivos líderes em torno da IA, em que as soluções são desenvolvidas ou adaptadas em estreita colaboração entre os líderes e os funcionários usuários da tecnologia, por meio de uma análise de toda a cadeia produtiva. Além disso, essas companhias tendem desenvolver soluções de IA internamente, em vez de comprar soluções de terceiros, como outras fazem (BALAKRISHNAN *et al.*, 2020).

Nelson e Winter (2002) mostram que, firmas maiores tendem a gastar mais em P&D e aumentar sua lucratividade a partir disso, o que faz com que elas cresçam e gastem mais com as inovações. Em relação às empresas com maior exposição à IA, existem mais indicadores de que elas estão ampliando sua vantagem competitiva do que evidências que as outras estão alcançando. Em consonância, outros resultados também indicam que esse grupo está obtendo ganhos superiores principalmente ao focalizar em aumento da receita com IA, ao invés da redução dos custos, embora também relatem isso. Nesse escopo de líderes analisadas, em uma pesquisa feita pelo Instituto McKinsey, o número dos relatos de ampliação da receita foi duas vezes maior que a diminuição de gastos (CHUI *et al.*, 2022).

Por gerarem mais resultados, naturalmente essas firmas estão preparadas para continuar gastando mais em esforços de IA do que outras. Como aplicam em uma porcentagem maior de receita, a taxa de crescimento do investimento também é superior, o que explica o aumento de vantagem crescente pelas empresas destaques. Outro aspecto é que isso tudo gera a elas um

ambiente mais atraente para talentos em IA, e mostra indícios de que essas organizações têm menos dificuldade em contratar para esses cargos. Mesmo as líderes sendo só 10% do total de firmas analisadas, elas contrataram duas vezes mais em áreas chaves como cientistas de dados de IA, engenheiros de *Machine Learning* e gerentes de produtos de IA. Dentro desses cargos mais qualificados, cerca de 70% das empresas relataram dificuldades de efetuarem contratações (CHUI *et al.*, 2022).

Qualificar ou requalificar são alternativas comuns à contratação, assim as firmas destaques estão viabilizando programas de IA e tecnologia com essas competências para aprimorar seus funcionários. Elas compreendem desde aprendizagem experimental, até cursos *on-line* e programas de certificação. Além disso, essas firmas são muito mais propensas a fornecer acesso a esses programas para capacitar os funcionários fora das áreas de TI. Com isso, elas estão melhor posicionadas para alcançar um sucesso sustentado pela IA. O aspecto positivo para as outras organizações é que, a partir dessas práticas, é deixado um modelo de êxito para se espelharem (CHUI *et al.*, 2022).

2.4 Riscos em relação à adoção de Inteligência Artificial

Em conjunto com os benefícios para o consumidor e valor para os negócios, a IA também dá origem a uma série de repercussões indesejadas. As mais evidentes, que incluem violações de privacidade, discriminação, acidentes e influência sob instituições, são preocupantes e devem incentivar a cautela. As consequências ainda não conhecidas ou que não foram possíveis de serem medidas são ainda mais preocupantes, que abrangem desde os dados alimentados nos sistemas de IA até a forma operacional dos algoritmos, a disrupções no mercado de trabalho e as interações entre humanos e máquinas (CHEATHAM & JAVANMARDIAN & SAMANDARI, 2019).

Repercussões desastrosas podem ocorrer quando um algoritmo de IA não funcionar como esperado. Já faz parte da trajetória da IA até a perda de vidas humanas, como o caso do carro autônomo da Uber que atingiu e matou um pedestre em Arizona, em 2018. Esses riscos se tornam desafios significativos para as organizações, desde danos à reputação e perdas de receita até ações regulatórias, investigações criminais e perda da confiança pública (CHEATHAM & JAVANMARDIAN & SAMANDARI, 2019).

Devido à difusão da IA ser muito recente, o nível de esforço necessário para identificar e controlar todos os principais riscos aparenta exceder as normas vigentes em muitas das corporações. Fazer progressos sólidos exige uma abordagem multidisciplinar envolvendo

líderes da alta gerência com o restante da companhia. A cobertura de riscos deve ocorrer em todas as áreas, desde TI e análise de dados, até o jurídico. Além disso, como é inevitável cometer erros em tecnologias recentes e ainda muito complexas, vigilância e fiscalização na empresa é essencial para evitar que problemas escalem (CHEATHAM & JAVANMARDIAN & SAMANDARI, 2019).

No âmbito destes riscos, a cibersegurança é relatada como um dos principais fatores para a maioria das empresas, pois crimes dessa natureza têm atingido a mídia, firmas e governos. Conforme citado no Relatório de Defesa Digital da Microsoft em 2022, o volume de ataques de senhas teve um aumento de 74% em apenas um ano, chegando a 921 ataques por segundo. Os ciberataques frequentemente têm impactos devastadores, gerando um custo médio de US\$ 4,35 milhões. Ao mesmo tempo, existe uma crise global de profissionais em cibersegurança, em que a demanda aumentou em média 35% no último ano, sendo que no Brasil chegou em 76%. O mercado não dispõe de profissionais qualificados suficientes para combater os ataques cibernéticos, colocando pessoas, empresas e governos ao redor do mundo em risco (BEHNCKEN, 2023).

Embora o uso de IA tenha acelerado, não houveram aumentos perceptíveis na redução de vários riscos relacionados a esta nas firmas não líderes. Já o motivo para não estarem mitigando os mesmos, é pela falta de capacidade ou entendimento em lidar com toda a gama de riscos que enfrentam. Em oposição, as corporações destaques em IA se envolvem em práticas de redução de riscos com mais frequência do que outras, como na verificação de algoritmos parciais, que relataram 2,6 mais vezes a mitigação desse risco na pesquisa feita pelo Instituto McKinsey citada anteriormente (CHUI *et al.*, 2022).

Nesse sentido, as empresas líderes também diferem em suas percepções sobre os perigos da IA. Analisando a relação oportunidade e risco, mais do que qualquer outro grupo, os pioneiros priorizam oportunidades ao invés do controle dos riscos, na busca de colher os benefícios antes dos concorrentes e novos entrantes. No entanto, como assegurar as vantagens econômicas pela IA é prioridade, o investimento para mitigar e controlar as consequências também são maiores (RANSBOTHAM *et al.*, 2018).

2.5 Expectativas do uso no futuro

Como já foi apontado antes, a implementação de IA em diversos setores da economia ocorreu de forma gradual durante as últimas décadas, sendo mais constante e intensiva nos últimos anos. Nesse sentido, essa adoção acontece de forma diferente tanto no mesmo setor

quanto em setores distintos, podendo variar dentre as empresas. Além disso, a intensidade desse movimento depende do nível de exposição e investimento do setor público e privado dos países em tecnologias e inovações.

Um caso de expectativa para o futuro da IA, que beneficia tanto a empresa quanto o meio ambiente, logo a sociedade em geral, é por meio da firma Vistra. A produtora de energia elétrica é uma das maiores nos EUA, operando em 12 estados com capacidade de mais de 39 gigawatts de eletricidade, o suficiente para fornecer energia para quase 20 milhões de residências. Em comparação, a maior produtora, Duke Energy, possui capacidade de 52 gigawatts. A Vistra se comprometeu a reduzir as emissões de poluentes em 60% até 2030, em relação a uma base de 2010, e alcançar zero emissões até 2050 (BOSWELL *et al.*, 2022).

Por meio de um algoritmo de IA desenvolvido durante dois anos, uma sede da Vistra operou com dois por cento a mais de eficiência após apenas três meses em operação, reduzindo por ano US\$ 4,5 milhões em gastos e 340 mil toneladas em emissões de carbono. Em comparação, relatos de companhias que constroem usinas de energia a gás investem milhões de dólares em P&D, e ao longo de quatro a cinco anos alcançam uma melhoria de apenas um por cento na eficiência (BOSWELL *et al.*, 2022).

Desde então, a Vistra expandiu a IA para outras 67 unidades de geração de energia em 26 usinas, obtendo em média um aumento de um por cento na eficiência e mais de US\$23 milhões em redução de custos. Com a expansão desse projeto, com mais de 400 modelos de IA, e outros já em desenvolvimento, foi possível reduzir cerca de 1,6 milhão de toneladas em emissões de carbono por ano. Com a expectativa da implantação dessa IA em toda a sua frota, para ajudar os operadores a tomar decisões ainda melhores, a empresa chega mais perto de alcançar sua meta em emissões (BOSWELL *et al.*, 2022).

Esse projeto reflete um traço fundamental da transformação da IA, que na Vistra e na maioria das firmas, o êxito da tecnologia não representa um simples acerto em uma tomada de decisão. O que ocorre são dezenas de melhorias significativas que se acumulam para fornecer um valor relevante, visando a acelerar um crescimento e tornar a tecnologia mais rentável que outras (BOSWELL *et al.*, 2022).

Como já expressado anteriormente, a intensidade da IA depende do nível tecnológico do país, sendo que os mais desenvolvidos naturalmente implementam mais IA, já países emergentes precisam de mais incentivos e investimentos para aumentar sua adoção à tecnologia. Em relação à América Latina, constata-se uma baixa difusão da tecnologia, que significa ainda muitas oportunidades para o uso dela. A Google, por exemplo, anunciou que vai

investir US\$ 1,2 bilhão ao longo dos próximos cinco anos em IA e outras tecnologias na América Latina (ARBEX, 2022).

Já em relação ao incentivo público, a Argentina, Brasil, Chile, Colômbia e Uruguai são os únicos países da região que lançaram estratégias nacionais autônomas de IA. Conselhos de especialistas e iniciativas políticas destacam a importância da IA para a transformação socioeconômica da região. Assim, as estratégias enfatizam como suas principais prioridades o cultivo de talentos locais e fortalecimento da infraestrutura tecnológica. O Brasil, por exemplo, propõe um programa nacional de alfabetização digital e revisão do currículo nacional de ensino do país para incluir cursos em IA e programação (THE ECONOMIST GROUP, 2022).

O fato é que, mesmo antes da crise gerada pela pandemia atingir a América Latina, a região sofria com o baixo crescimento econômico e a distribuição desigual de ganhos. Devido a muitos setores no território atuarem com uma baixa eficiência quando comparado com outros mercados mais desenvolvidos, os ganhos de produtividade com IA são cruciais para melhorar a performance econômica do território. Estimativas do Instituto McKinsey indicam que, mesmo estando defasada em relação a outros na adoção de IA, o valor a ser liberado pelo uso dessa tecnologia na América Latina é de cerca de US\$ 600 bilhões a US\$ 1 trilhão anualmente em ganhos de produtividade (CAFFERATA *et al.*, 2020).

Mesmo em crise, a pandemia causou um aumento sem precedentes no investimento privado em tecnologia na América Latina, com aplicação de capital de risco aumentando mais de três vezes de 2020 para o ano seguinte. Desde 2019, mais de US\$ 20 bilhões de financiamento de capital de risco foram destinados a *startups* de TI. Historicamente, a região era reconhecida por receber pouco investimento em tecnologia, mas com o deslocamento das atividades *on-line* pela pandemia, gerou-se um senso de que a esta precisava recuperar o atraso (THE ECONOMIST GROUP, 2022).

Outro fator que contribuiu foi a forma como as maiores indústrias da América Latina – finanças, saúde e infraestrutura – eram dominadas por apenas algumas empresas. Um exemplo particular é o setor financeiro do Brasil, do qual 80% é controlado por apenas cinco bancos. A insatisfação pública com as ineficiências e modelos de negócios arcaicos criou a tempestade perfeita para que essas firmas fossem contestadas. Assim, modelos de negócio muito mais voltados para a tecnologia, como o Nubank, que é uma plataforma de serviços financeiros digitais e que usa IA para decisões de crédito, ganham força no mercado e estimulam a concorrência a elevar a qualidade e inserção tecnológica nos seus produtos e serviços (THE ECONOMIST GROUP, 2022).

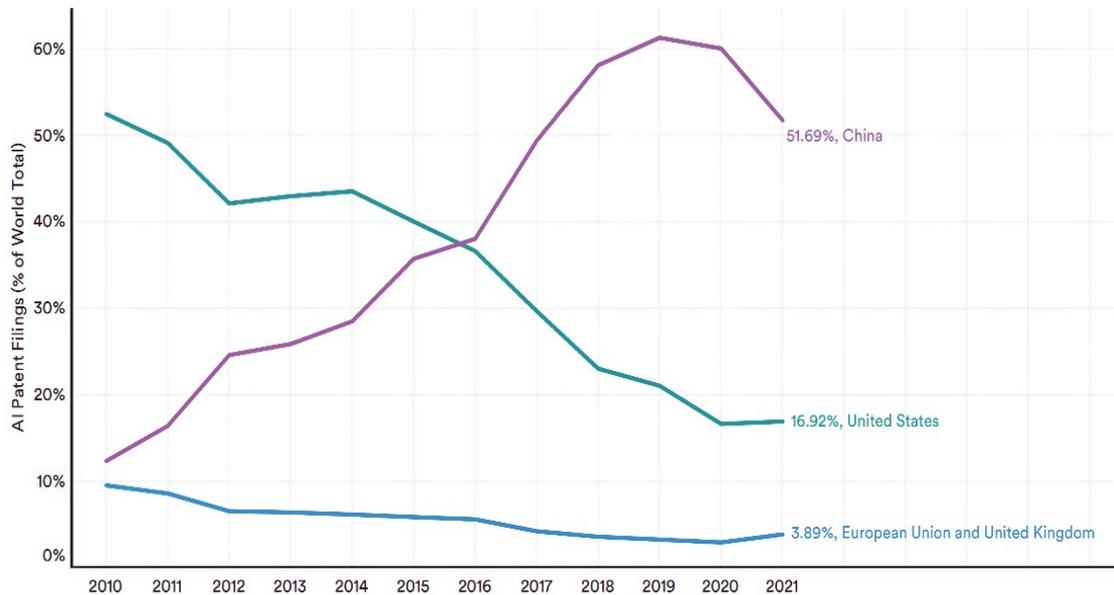
Com isso, muitos investimentos em IA devem ser feitos na América Latina. Um relatório do The Economist aponta um estudo que estima um crescimento de 38% no mercado de IA em aplicações de saúde entre 2019 e 2027. Esse setor se tornou um foco em investimentos, impulsionado pela necessidade de aliviar os sistemas de saúde da região, que estão sobrecarregados pela demanda dos pacientes. Embora a agricultura seja uma das maiores indústrias da região, o volume de investimentos em tecnologia nela continua relativamente baixo. No entanto, o futuro tende a ser promissor, o número de *startups* nessa área aumentou de 64 para 457 entre 2005 e 2018. Já o setor financeiro recebeu a maior parte dos investimentos, com cerca de 40% de todos os investimentos destinados às empresas de financeiras tecnológicas (THE ECONOMIST GROUP, 2022; VITON & LESKA & TEIXEIRA, 2019).

Os maiores desafios políticos enfrentados pela América Latina são a fragmentação da infraestrutura e a urgência de promover o desenvolvimento de talentos locais. Indivíduos altamente capacitados frequentemente saem da região em direção a outras localidades do globo, especialmente para a América do Norte, com o intuito de se integrar a empresas de tecnologia mais competitivas. Nesse sentido, as duas potências disparadas em IA são a China e EUA, cada uma com suas individualidades (THE ECONOMIST GROUP, 2022).

Como um dos líderes em IA, a China produziu cerca de um terço dos artigos científicos e citações de IA em todo o mundo em 2021. Em termos de investimento, representou quase um quinto do financiamento global de investimento privado em 2021, atraindo US\$ 17 bilhões para *startups* de IA. A maioria das aplicações de IA adotadas até agora foi em indústrias voltadas ao consumidor, impulsionadas pela sua base de consumidores de internet, que é a maior do mundo. A partir desses investimentos, estima-se que, em setores como manufatura, automotivo, transporte e logística, a IA poderá gerar mais de US\$600 bilhões em receita anualmente para o país (SHEN *et al.*, 2022).

Um relatório da Universidade de Stanford mostra o rápido crescimento das solicitações de patentes de IA nas principais regiões líderes na tecnologia. Como se verifica no gráfico 1, em 2010, as solicitações de patentes da China representavam apenas cerca de 12% do total mundial e subiram para 52% em 2021. Em comparação, os Estados Unidos passaram de 52% em 2010 para 17% em 2021, e a União Europeia e Reino Unido diminuíram de 10% em 2010 para 4% em 2021 (LI, 2022).

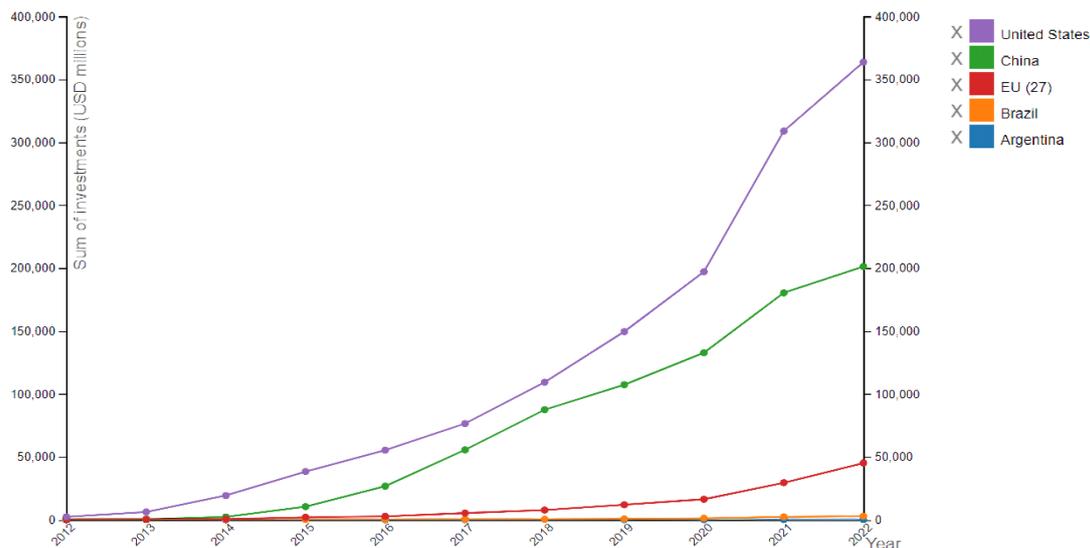
Gráfico 1 – Pedidos de patentes de IA por área geográfica, 2010 – 2021 (% total do mundo)



Fonte: MASLEJ, N. *et al* (2023)

Como apresentado pelo gráfico 2, em termos de investimento privado em IA, os EUA apresentam números muito maiores que de outros países, incluindo a China que também é modelo em IA, diferença que chega a quase o dobro na última década. Já a Europa como terceiro maior número de investimentos se encontra muito atrás, sendo que outras nações, como alguns que adotaram estratégias de IA da América Latina, apresentam números insignificantes perante os líderes. Vale notar que, foi relatado que dentre esses investimentos privados em IA nos EUA, parte deles possui origens chinesas, de acordo com o FBI e pesquisadores (LI, 2022).

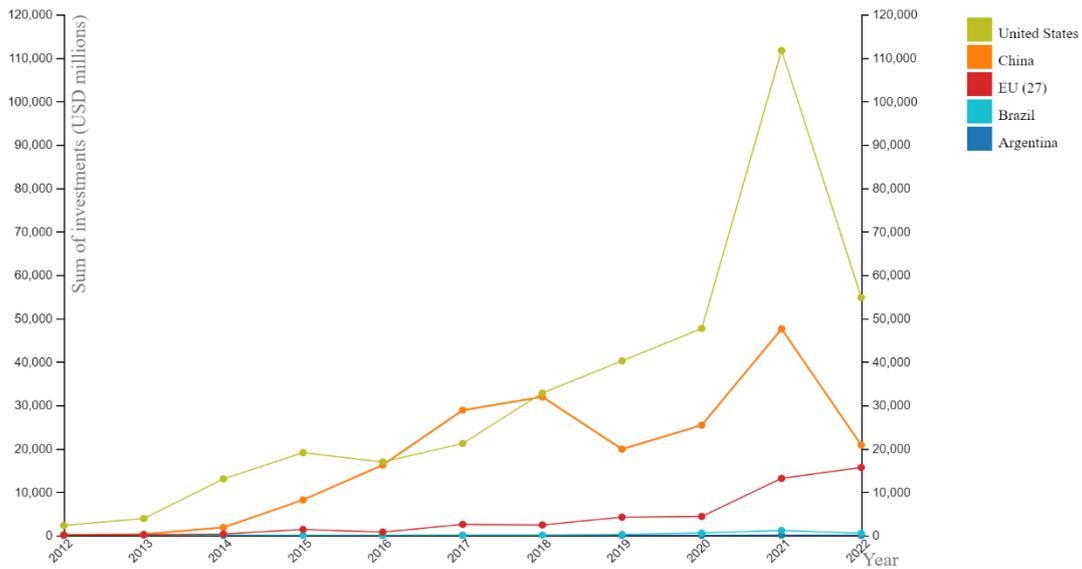
Gráfico 2 – Somatório de investimentos privados em IA de 2012 – 2022 (US\$ milhões)



Fonte: OECD (2023b)

Por outro lado, em termos de valor absoluto dos investimentos (Gráfico 3), devido à queda dos EUA e China nesse quesito, pela primeira vez na última década, o valor do capital privado aplicado globalmente em IA ano após ano diminuiu. Em contraste, países como Brasil e Argentina não possuem valores significativos em comparação com os líderes. Além disso, o gráfico 3 mostra que enquanto os investimentos dos EUA e China começaram a crescer em 2014, a UE somente em 2020 começa a incentivar neste mercado, sendo o terceiro maior atualmente, que enfatiza a disparidade entre os outros países em relação à IA. O investimento privado global em IA foi de US\$ 220 bilhões em 2021, e passou a US\$ 122 bilhões em 2022, o que representou um retrocesso de em média 45%. O número total de eventos de financiamento relacionados à IA, bem como o número de empresas de IA recém-financiadas, também reduziu, resultado da queda de capital no mercado, causado pela alta global dos juros (OECD, 2023b; MASLEJ *et al.*, 2023).

Gráfico 3 – Investimento privados em IA de 2012 – 2022 (US\$ milhões)

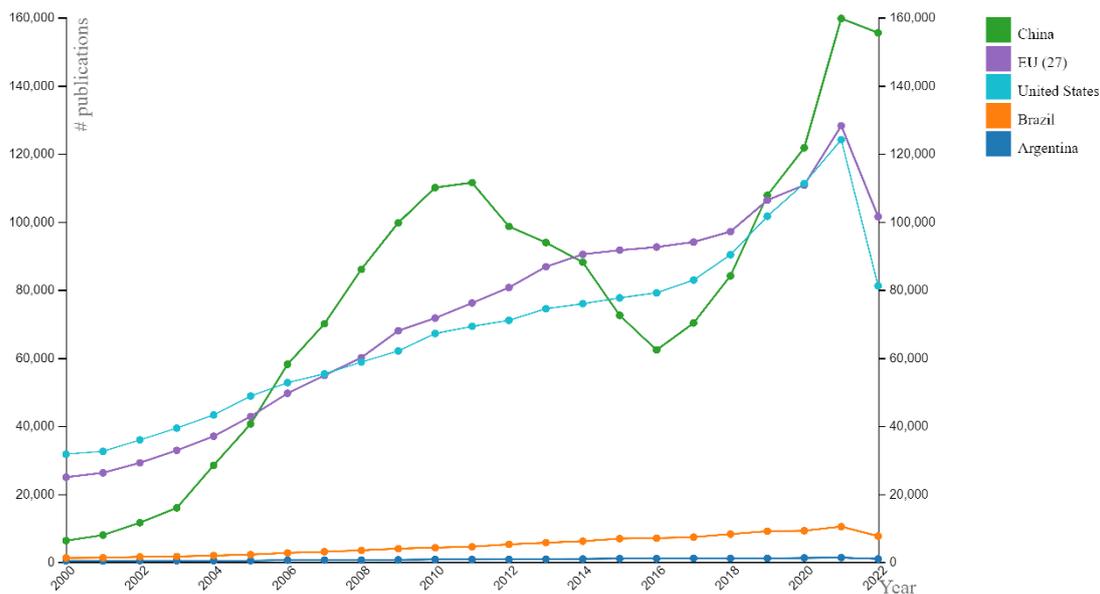


Fonte: OECD (2023b)

Na China, é evidente que o governo lidera o desenvolvimento de estratégias e projetos de IA no país, enquanto isso, o setor privado, por mais que represente uma grande parte dos investimentos, têm pouca influência no plano chinês. Isso implica que o setor público pode identificar rapidamente as melhores perspectivas para a IA, e direcionar financiamento substancial para os projetos de seu interesse e que tenham maior grau de sucesso (SHEN *et al.*, 2022).

Em contrapartida, o plano dos EUA é muito mais colaborativo, com o governo federal, setor privado e universidades desempenhando papéis ativos, mediante suas próprias convicções. No momento, o foco parece ser cultivar uma ampla variedade de projetos em vez de se concentrar em um conjunto mais estreito, mantendo o livre-arbítrio e liberdade criativa dos desenvolvedores e difusores da IA. O governo tem afirmado repetidamente que regulamentações e regras estatais consideradas excessivas são frequentemente vistas como barreiras à inovação (SHEN *et al.*, 2022).

Gráfico 4 – Número de publicações relacionadas a IA por país de 2000 – 2022



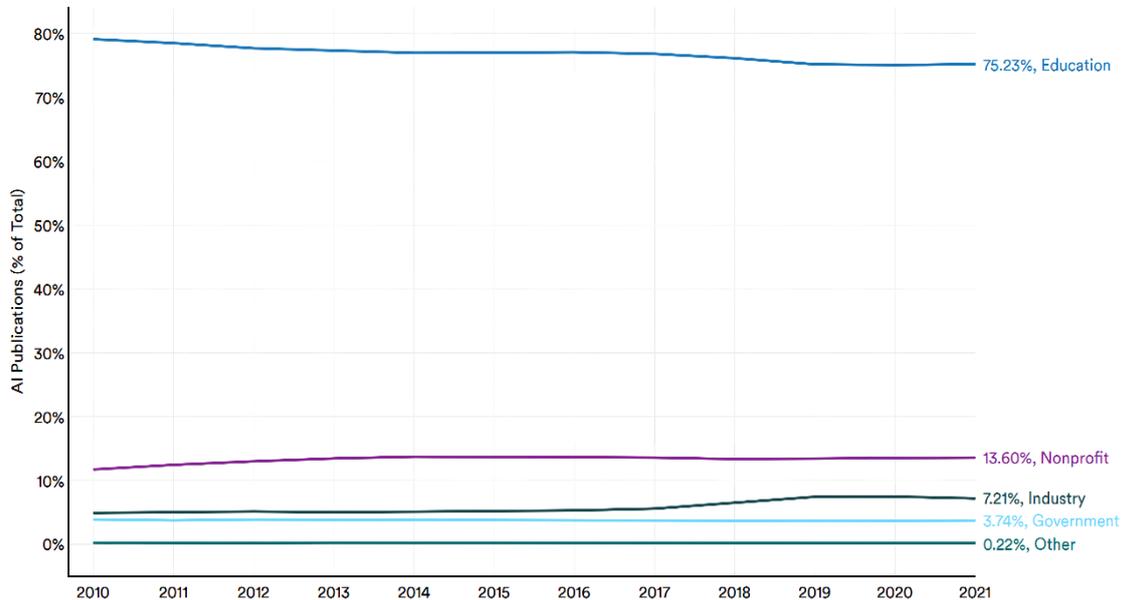
Fonte: OECD (2023b)

Por se destacar no cenário global de pesquisa, desenvolvimento e formação de políticas em IA, a China se tornou uma líder nesse meio, possuindo o maior número de publicações sobre a tecnologia no mundo. Como pode ser visto no gráfico 4, atualmente ela lidera, chegando a quase 160 mil publicações, quase o dobro dos EUA. A UE e os EUA têm uma trajetória parecida em relação às publicações. Assim, pelos gráficos 3 e 4, pode-se concluir que existe um claro atraso na adoção e incentivo à IA nas nações europeias, uma vez que as pesquisas chegam muito antes que o investimento na região. Já representando a América latina, por serem os mais notórios em IA na região, o Brasil e Argentina ainda apresentam condutas muito distintas dos líderes em IA. Ao longo de mais de duas décadas, a China fez-se profundamente envolvida na rede internacional de P&D em IA, sendo coautora de artigos com agentes no exterior, hospedando laboratórios corporativos de IA americanos e ajudando a expandir as fronteiras da pesquisa global em IA (KERRY & MELTZER & SHEEHAN, 2023).

Durante parte desse período, essas conexões e suas implicações foram amplamente ignoradas no mundo político, sendo ditadas pelos pesquisadores, universidades e corporações que as arquitetavam. No entanto, nos últimos cinco anos, esses laços entre a China e as redes globais de P&D têm enfraquecido e tem sido cada vez mais reavaliado por governos, universidades e empresas, comprometendo o futuro da IA colaborativa entre uma das líderes na tecnologia e o resto do mundo. Esse movimento leva a IA a um caminho ainda mais

individualista e excludente da tecnologia, que condiz com o plano de interesse da China e com suas ações (KERRY & MELTZER & SHEEHAN, 2023).

Gráfico 5 – Publicações relacionadas a IA por setor, 2010 – 2021 (% total do mundo)



Fonte: MASLEJ, N. *et al* (2023)

Como apresentado no gráfico 5, os setores responsáveis pelas publicações de IA, podem ser filiadas à educação, o governo, a indústria, ser sem fins lucrativos e outros setores, sendo que o setor de educação domina em todas as regiões. Assim, como a maioria dos estudos são vinculados às instituições de educação, e não ao governo, isso pode explicar o porquê de, mesmo com constantes disputas nas últimas décadas, somente nos anos recentes a colaboração internacional com a China ter caído. Um outro ponto a se levantar é que, em 2018, quando começam a cair as colaborações, também foi a data de início da guerra comercial entre China e EUA (MASLEJ *et al.*, 2023).

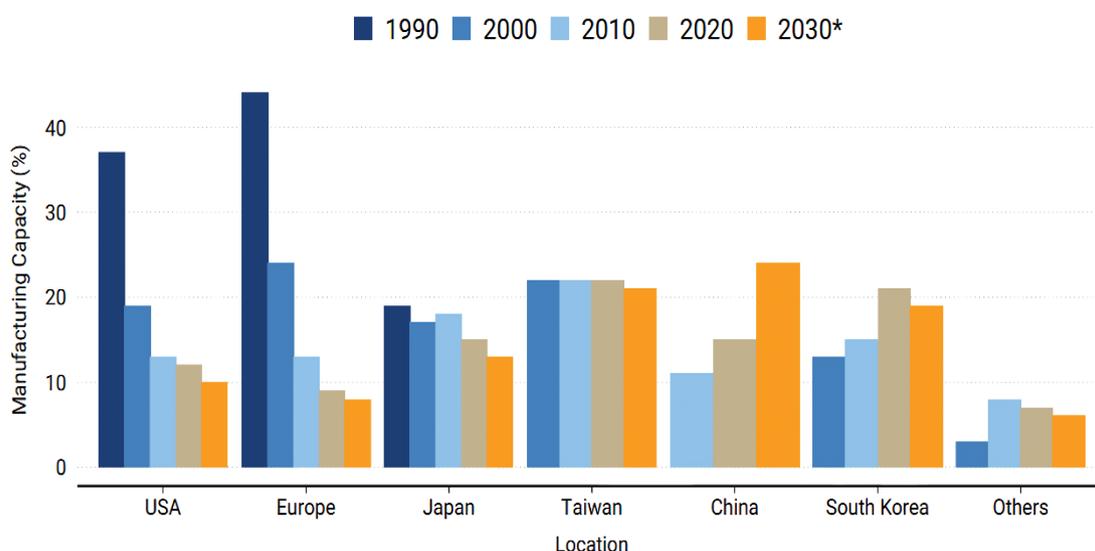
As principais justificativas relatadas pelos EUA e EU para o distanciamento em relação a IA chinesa são: os riscos gerados pelo aumento da capacidade da IA em si e seus impactos tanto na competitividade econômica e segurança; as políticas fortalecedoras dessas capacidades, que incluem investimentos direcionados pelo Estado e transferências pelo exterior de legitimidade desconhecidas, sem nenhum tipo de transparência; o aumento das ambições excludentes do governo chinês, que faz dela uma concorrente direta do EUA; e o uso antiético de IA pela China, notoriamente na utilização para vigilância em massa de seus cidadãos (KERRY & MELTZER & SHEEHAN, 2023).

Dessa forma, em uma conferência recente do Conselho de Comércio e Tecnologia entre a União Europeia e os Estados Unidos, ambos declararam sua oposição a modelos de IA que não respeitam as leis e os direitos humanos, como os sistemas de pontuação social. O alvo implícito das críticas foi o sistema de "crédito social" da China, que utiliza uma ampla variedade de dados para avaliar a pontuação de uma pessoa e determinar as suas permissões na sociedade, como a compra de passagens aéreas ou de trem. A crítica é que o uso desse programa pelos governos autoritários implica em um controle social em larga escala, e representam ameaça à liberdade e direito dos cidadãos, incluindo a censura de discursos e outras atividades de expressão, além de reforçar os sistemas de vigilância arbitrária ou ilegal (LARSEN, 2022).

Nesse sentido, a principal desavença no momento é a disputa entre os EUA e a China, ambos líderes disparados em IA. Nos EUA, um relatório de 2021 da Comissão de Segurança Nacional de Inteligência Artificial do país (NSCAI) definiu a "corrida da IA" como uma competição baseada em valores na qual a China deve ser vista como um concorrente direto. No relatório, o NSCAI foi ainda mais longe e recomendou a criação de chamados "pontos de estrangulamento" que limitem o acesso da China aos semicondutores americanos para retardar o progresso deles em algumas áreas de desenvolvimento tecnológico (LARSEN, 2022).

Assim, medidas posteriores a essa disputa já foram tomadas, em agosto de 2022, o Departamento de Comércio dos EUA proibiu a Nvidia, importante corporação especializada no desenvolvimento de componentes eletrônicos, de vender suas unidades de processamento de gráficos (GPUs) A100, A100X e H100 para clientes na China. Essa medida tem como finalidade retardar o progresso da China no desenvolvimento de semicondutores, além de impedir que chips avançados sejam usados para aplicações militares. A medida teve como justificativa "manter tecnologias avançadas fora das mãos erradas", mesmo com a Nvidia expressando que resultará em sérias consequências para suas vendas globais de semicondutores (LARSEN, 2022).

Gráfico 6 – Produção global de semicondutores por país, 1990 – 2030 (% total do mundo)



Fonte: BRANDIT, J. *et al* (2022)

Embora os EUA e a UE tenham lançado recentemente novas iniciativas e políticas industriais voltadas explicitamente para semicondutores, a China tem há muito tempo cultivado essa indústria. Essa tecnologia tem relação direta com a capacitação de modelos de IA, que é de extrema significância para o governo chinês, assim, já em 2014, foi criado o Fundo de Investimento da Indústria Nacional de Circuitos Integrados. Isso corresponde com os ideais mostrados na estratégia IA da China, que no intuito de ser líder internacional em IA até 2030, pretende comandar também todos os segmentos da cadeia de suprimentos de chips. Desse modo, o gráfico 6 mostra essa trajetória, em que os EUA e UE perdem a hegemonia que tinham em 1990, já a China ganha relevância e pelas previsões vão conseguir assumir a posição almejada de principal produtor do setor (LARSEN, 2022).

Nesse cenário de disputas entre nações para garantir vantagens competitivas pela adoção de IA, o próximo capítulo retrata de que formas o uso ético, ou não, da IA pode influenciar nas relações e nas medidas protecionistas tomadas por cada país. Além de que, a tecnologia pode representar um problema para os próprios países, a sua complexidade e possibilidade de uso abrangente permitem várias brechas para erros e oportunismos. Assim, serão apontados tanto os potenciais problemas como as regulações vigentes e previstas para combater os mesmos.

CAPÍTULO 3: DESAFIOS ÉTICOS E MORAIS DA INTELIGÊNCIA ARTIFICIAL

3.1 Relevância da atuação governamental em benefício da sociedade

À medida que cresce o uso e eficiência da IA, mais complexo fica de garantir que ela exerça apenas os fins realmente estimados e não gere consequências indesejadas, transgredindo barreiras éticas e legais. Além do sistema alcançar os objetivos de quem o utiliza, deve-se considerar a possibilidade dele impor externalidades sobre os outros, logo existe um alinhamento social entre seus efeitos e o grupo/sociedade que ele afeta. Os atritos geralmente surgem de conflitos entre objetivos individuais e sociais, o que eleva a importância da governança de IA para mediar tais conflitos. Isso gera tanto a necessidade de mudança de normas já existentes, quanto a criação de novas medidas que acompanhem o nível mutável de desenvolvimento da IA (KORINEK & BALWIT, 2022).

Dessa forma, pode-se distinguir esse alinhamento pela forma direta, quando um sistema de IA busca objetivos condizentes com os de seu operador, independentemente de impor externalidades a outras partes. Já pelo social a IA tem seus fins mais consistentes com os da sociedade, levando em consideração o bem-estar de todos que são impactados. Uma vez que o alinhamento social ideal é inatingível na prática, é viável determinar uma ordenação parcial das preferências populares que atendam o maior número de pessoas, e coordenar a IA com esse conjunto mais limitado de instruções. Essa medida tende a exigir intervenções externas e um quadro de governança mais amplo (KORINEK & BALWIT & 2022).

O problema de alinhamento social pode se manifestar em diferentes escalas, desde pequenos subgrupos da sociedade até esferas maiores, sendo a característica comum entre eles que em todos casos alguém é afetado por externalidades. Em geral, o grau de concordância dentro de um grupo diminui à medida que o tamanho do grupo aumenta, com isso, comunidades menores devem chegar a um acordo sobre os resultados desejáveis da IA com mais facilidade do que grupos maiores de pessoas. Isso implica que as preferências dos grupos levarão a ordenações mais parciais à medida que o tamanho do grupo aumenta (KORINEK & BALWIT, 2022).

Independentemente da esfera, alcançar o alinhamento coletivo de uma IA tende a ser mais desafiador do que atingir o seu propósito direto. Enquanto o objetivo do criador de um sistema de IA é resolver o problema para que o sistema busque os resultados desejados, a orientação social, ou sua falta, pode ser uma reflexão tardia do mesmo. A concordância entre ambos os lados seria certa somente se, o operador fosse perfeitamente altruísta e

internalizasse todas as externalidades que a IA impõe aos outros de maneira ética. No entanto, o operador pode não se importar em causar danos aos outros, desde que alcance seus resultados, ou podem surgir males sem que ele sequer esteja ciente disso, e ele não ter incentivos suficientes para descobri-los (KORINEK & BALWIT, 2022).

À medida que cientistas e pesquisadores desenvolvem métodos para analisar os riscos e benefícios da IA, as colaborações junto com o governo, políticos e agências nacionais resultarão na elaboração de medidas mais fiéis à realidade e baseadas em evidências, para melhor regular a IA. Assim, a introdução de padrões necessários para tornar a IA mais confiável afetaria como as empresas constroem e implementam esses sistemas. Ao mesmo tempo, tornar essa tecnologia mais transparente facilitaria a prevenção de danos causados ao expor o impacto da IA para as partes afetadas, diretas e indiretas, aumentando a conscientização sobre como a IA transforma vários aspectos da vida individual e da sociedade, por meio de abordagens mais éticas e humanísticas (CALISKAN & STEED, 2022).

No entanto, mesmo que os estudos desempenhem um papel importante nessa ajuda ao setor público, eles frequentemente promovem algum interesse privado específico. Entretanto, os estudos são esperados para apresentar argumentos que persuadem racionalmente as pessoas de que uma política é melhor do que outra, em termos de valores amplamente aceitos e que são considerados aplicáveis à sociedade como um todo, em vez de a um grupo específico (NELSON & WINTER, 1982).

Além disso, a realidade é um pouco mais desafiadora para os criadores de leis do que o esperado, por exemplo, quando o processo de criação de uma IA é feito por colaboração entre várias firmas, e por *Machine Learning* ela desenvolve por conta própria padrões não éticos de comportamento, como exclusão de minorias. Nesse caso, os reguladores terão que decidir quem é responsável pelo erro, podendo transgredir barreiras nacionais com legislações diferentes. O cenário de pesquisa em IA altamente difuso e internacional significa que, muitas decisões sobre cooperação em projetos de pesquisa específicos cabem a pesquisadores individuais e instituições. No entanto, os governos têm um papel a desempenhar nesse processo, fornecendo aos pesquisadores e instituições as informações e o apoio necessário para agir de forma responsável (KERRY & MELTZER & SHEEHAN, 2023).

Nesse sentido, os governos enfrentam inúmeras barreiras, incluindo a falta de profissionais altamente capacitados e especializados, recursos limitados para investir em pesquisa e inovação em IA, e regulamentações pouco claras em relação à ética, segurança e transparência nesse meio. Isso pode não só dificultar a adoção do uso de IA pelos governos,

como não gerar interesse do setor privado, principalmente em países subdesenvolvidos onde já é baixo. Por isso, ao desenvolver e implantar a IA, é crucial que os governos considerem e abordem proativamente o cenário de rápidas mudanças dessa tecnologia, os riscos desde éticos até de segurança às quais a IA podem expô-los são altos, porém a abstenção de seu uso pode ser tão prejudicial quanto (BERGLIND & FADIA & ISHERWOOD, 2022).

Os governos podem potencialmente desempenhar um papel ativo e crítico na captura dos muitos benefícios da implementação de IA, variando de acordo com as estruturas políticas e econômicas do país. Pode ser viável a criação de um setor ou órgão responsável por iniciativas de IA, composto por profissionais qualificados, institutos de pesquisa, *startups* e grandes empresas. Alguns países também lançam programas de assistência para IA onde geram várias oportunidades. Cingapura, por exemplo, conta com projetos de cidades e mobilidade urbana inteligente. Além desse caso, muitas nações criaram ambientes centralizadas em IA para gerar valor a partir desta, sendo responsável por originar iniciativas que impulsionam, ou regulações que restringem a abordagem de IA no país (BERGLIND & FADIA & ISHERWOOD, 2022).

No entanto, como já apontado antes, as rápidas transformações da IA dificultam para os governos e outros órgãos reguladores acompanharem o ritmo, e implementarem políticas e estruturas de governança efetivas voltadas para aproveitar o potencial e mitigar os riscos da IA. Devido ao potencial que possuem, tecnologias como essa, de aplicação generalizada e global, não podem ser realizadas e governadas por um único país ou ator econômico. A infraestrutura computacional, habilidades e dados necessários para desenvolver e implantar um único sistema de IA muitas vezes atravessam fronteiras jurídicas. Portanto, é necessária cooperação e coordenação internacional, para garantir que as inovações em IA melhorem o bem-estar das pessoas em áreas como educação, segurança pública e saúde, sem comprometer a jurisdição e qualidade de vida de outros países (PERSET *et al.*, 2022).

3.2 Incertezas sobre a Inteligência Artificial

A IA assim como outras tecnologias pode ser usada de forma ética ou não, como ferramenta cabe inicialmente ao usuário definir qual será a utilidade da mesma, o que gera certo nível de incerteza a esse tipo de tecnologia. Para piorar a situação, a IA conta com duas características que potencializam essas dúvidas: o alto nível de disseminação global, logo afeta incontáveis vidas em várias áreas, e a capacidade de aprender por contra própria por *Machine Learning*, ou seja, está sujeita a eventualidades. Um exemplo é a IA de reconhecimento facial na China, o governo tem desempenhado um papel central e crescente na adoção desse tipo de

tecnologias para supervisionar espaços públicos, onde 626 milhões de câmeras de monitoramento foram instaladas no país até 2020 (LARSEN, 2022).

Nesse sentido, a crescente sofisticação da IA em tecnologias de vigilância digital têm levantado preocupações em relação à privacidade e aos direitos humanos, sendo a China uma das líderes no desenvolvimento e uso dessas tecnologias tanto para uso doméstico quanto internacional. No entanto, outras nações também estão ativas nesse espaço, incluindo os EUA, Israel, Rússia, Japão, Coreia do Sul e diversos países europeus. Empresas dos Estados Unidos, em particular, desempenham um papel importante ao fornecer o *hardware* subjacente para as tecnologias de vigilância (PETERSON & HOFFMAN, 2022).

Essas tecnologias de IA e vigilância digital são utilizadas em uma série de cenários, como apoio do trabalho policial ao ajudar a otimizar as operações e resolver problemas cotidianos, melhorando as capacidades de previsão e resposta a emergências. No entanto, essas tecnologias também podem ir contra os direitos e interesses das pessoas que elas afirmam proteger, reafirmando os preconceitos e danos que elas deveriam eliminar. Alguns dos casos mais graves de uso incluem espionagem de figuras políticas e repressão das populações muçulmanas uigures e turcas na China. Assim, as incertezas no uso da IA giram em torno das decisões tomadas pelos seus usuários e regulações autônomas feitas pelos próprios países (PETERSON & HOFFMAN, 2022).

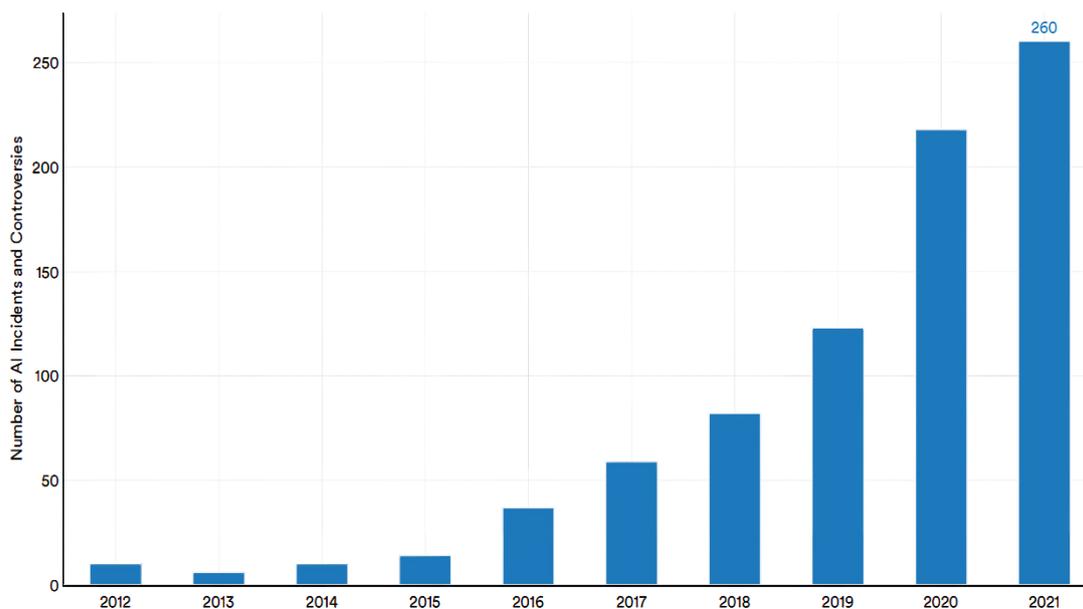
Outro exemplo de tipo de risco ocorre no ambiente totalmente virtual, é o surgimento de um novo tipo de empresa de mídia, as redes sociais, que usam de conteúdo gerado pelos usuários e disseminam essas informações por meio de um algoritmo de IA. Em vez de jornalistas ou mediadores, essas companhias usam sistemas de recomendação para moderar o conteúdo e decidir qual informação servir aos usuários. O objetivo desse método é maximizar o engajamento, por meio de indicações que sejam mais prováveis de assemelhar a gosto do usuário e prender sua atenção. Dessa forma, é possível maximizar o tempo que a pessoa passa na plataforma e interage com o conteúdo, que está diretamente relacionado com a receita gerada pela empresa com anúncios (BULLOCK & KORINEK, 2022).

O conteúdo disseminado nas redes sociais por usuários comuns, anunciantes e também por firmas de mídia tradicionais muitas vezes não é verificado e/ou nem mesmo tem intuito de ser informativo. No entanto, ainda são necessárias decisões sobre qual conteúdo cada usuário recebe, ou seja, o conteúdo ainda requer algum mecanismo de coordenação e moderação para distribuir as informações. Como consequência, uma série de resultados indesejáveis têm provocado enormes externalidades negativas à sociedade, desde o aumento da polarização

política e propagação de desinformação, até a disponibilização de conteúdos tóxicos para menores de idade (BULLOCK & KORINEK, 2022).

Nesse sentido, o risco não só está ligado à ética da empresa no uso da sua IA de recomendações, mas no modo que esse programa aprende de forma autônoma, por conta própria toma decisões e gera eventualidades desfavoráveis para os usuários ou até mesmo à companhia. Para combater isso, as redes sociais têm contratado moderadores de conteúdo para marcar, remover ou bloquear conteúdos considerados inconvenientes, que sejam violentos, sexualmente explícitos ou ofensivos. Mas a remoção de postagens expõe as companhias à crítica de censura e não resolve o maior problema: enquanto um determinado conteúdo não chega a atenção dos moderadores humanos, eles são considerados adequados e ficam livres para serem propagados pelo sistema de recomendação (BULLOCK & KORINEK, 2022).

Gráfico 7 – Número de incidentes e controvérsias envolvendo IA de 2012 – 2021



Fonte: MASLEJ, N. *et al* (2023)

O Repositório de Incidentes e Controvérsias de IA, Algoritmos e Automação (AIAAIC) é um banco de dados aberto, independente e público de incidentes e controvérsias relacionados à IA. Lançado em 2019 como um projeto privado para analisar os riscos da IA, ele evoluiu para uma iniciativa abrangente que rastreia as questões éticas associadas à tecnologia de IA. Como pode ser verificado no gráfico 7, houve um aumento progressivo de incidentes com IA relatados na última década, isso ocorre por consequência do grau crescente de integração da IA na sociedade, e da maior conscientização das pessoas em relação as maneiras pelas quais a IA pode

causar danos. No entanto, à medida que esse esclarecimento cresce, o rastreamento de incidentes também acompanha, sugerindo que incidentes mais antigos podem estar subnotificados. Um exemplo de incidente notório foi, em março de 2022, que um vídeo circulou nas redes sociais e em um *site* de notícias da Ucrânia, supostamente mostrava o presidente ucraniano direcionando seu exército a se render na luta contra a Rússia. Porém, foi revelado posteriormente que o vídeo era falso, criado por uma IA (MASLEJ *et al.*, 2023).

3.3 Danos éticos e legais gerados pela Inteligência Artificial

Em algumas instâncias, o que era considerado apenas um risco em relação à IA, pode se tornar claramente uma violação legal ou ética gerada pela tecnologia. Assim, diferentemente dos exemplos anteriores, nesses segmentos as situações de fato geraram danos a um indivíduo ou sociedade. Um exemplo é quando as aplicações de IA não tratam todos os grupos de forma igual, devido ao processo de *Machine Learning*, o auto aprendizado da IA se baseia em dados previamente existentes. Logo, o algoritmo pode aprender e reforçar preconceitos presentes na sociedade, que colocam certos grupos em desvantagem. Essa parcialidade negativa presente no processo tem sido identificada em diversos contextos, como sentenças judiciais, admissões educacionais e decisões de contratação (DAVENPORT & KATYAL, 2019).

Mesmo quando os criadores do algoritmo não têm a intenção de gerar parcialidade ou discriminação, eles devem ser responsáveis por identificar e prevenir tais problemas, corrigindo-os quando descobertos. É claro que a desigualdade não é um problema novo, pessoas e empresas que utilizam processos de tomada de decisão tradicionais também cometem imparcialidades e erros de julgamento, possivelmente mais que a IA. No entanto, devido à capacidade e velocidade computacional na aplicação dos modelos e processamento dos dados, esses danos podem rapidamente ganhar uma enorme dimensão antes de serem descobertos (DAVENPORT & KATYAL, 2019).

Em um caso, um algoritmo de recrutamento utilizado pela Amazon gerava imparcialidade contra mulheres. Isso foi decorrência de o sistema de IA estar aprendendo por meio de dados submetidos por candidatos predominantemente masculinos, ao longo de um período de dez anos. Como resultado, o algoritmo concluiu por conta própria que os candidatos masculinos eram preferíveis, e assim que o problema foi descoberto, a Amazon encerrou a utilização dessa IA (THEOTO *et al.*, 2022).

Já em outra situação, a Comissão de Igualdade de Oportunidades de Emprego (EEOC) nos EUA, divulgou informações técnicas sobre como alguns algoritmos de contratação eram

discriminatórios contra pessoas com deficiências. Os dados usados para treinar a IA de contratação não representam adequadamente as diferentes deficiências dos candidatos, e pode ter sido projetada sem mesmo levar em consideração pessoas com deficiências. Como a tecnologia aprende identificando padrões e replicando-os, essas ferramentas podem reforçar as desigualdades existentes na sociedade, refletindo os preconceitos existentes no mundo real, e passando despercebidos devido à falta de transparência dos modelos de IA (ENGLER, 2022b).

A capacidade de integrar facilmente a tecnologia de IA em várias áreas da economia levou ao surgimento de IAs imperceptíveis. Ao contrário de carros autônomos ou assistentes pessoais, sistemas que utilizam a IA invisível não anunciam esse fato, em vez disso, são usados em segundo plano para prever resultados e tomar decisões (FONSEKA, 2017).

Nesse sentido, muitas vezes é difícil ou impossível saber por que um algoritmo tomou uma decisão, pois atuam como uma "caixa preta", não sendo possível determinar como seu resultado foi produzido, além de serem considerados segredos comerciais das empresas privadas que o desenvolvem e comercializam. O problema é potencializado ainda mais por não ser possível interpretar ou explicar nem mesmo a lógica do modelo. Embora a plena transparência dos modelos seja benéfica para sociedade, os líderes que têm sua tecnologia de IA como um ativo competitivo resistirão em compartilhá-los (DAVENPORT, KATYAL, 2019).

No meio desses processos opacos, ocorre a definição parcial de preços, quando valores distintos são mostrados a diferentes consumidores para um mesmo produto. Alguns *sites* de reserva de hotéis e viagens usam um algoritmo que determina quanto cobrar dos consumidores com base em seus dados pessoais. Assim, mostram preços mais baixos aos usuários que navegam com um dispositivo de valor mais barato ou mostram preços mais altos aos consumidores que já alugaram quartos de hotel mais caros no passado. Considerando que esses dados foram encontrados de forma legal, seria um exemplo de uso antiético de IA, porém, esses processos não são transparentes para o público, que leva ao questionamento se realmente ocorre de forma legal (FONSEKA, 2017).

No campo da educação, instituições de ensino superior nos EUA estão usando algoritmos que alocam estrategicamente bolsas de estudo para convencer mais estudantes a se matricularem. Por meio dessa IA, as faculdades variam o custo de matrícula com base na disposição dos estudantes a pagar, reduzem a quantidade de financiamento de bolsas de estudo oferecido aos estudantes e abrem a possibilidades de discriminação por meio da tecnologia. A educação superior nos EUA já enfrenta baixas taxas de graduação, alto endividamento

estudantil e desigualdade para minorias raciais – crises que os algoritmos de matrícula tendem a ajudar a perpetuar (ENGLER, 2021).

Nesse sentido, pelo fato da IA indicar as pessoas mais prováveis, uma universidade teve um aumento de 23,3% na taxa de matrícula de estudantes de fora do estado, além de um aumento de 10% no valor da matrícula por meio dessa otimização de bolsas de estudo. Em escala, isso pode resultar em milhões de dólares em receita adicional de mensalidades, que é a razão pela qual tantas faculdades estão adotando esses algoritmos. Como benefício adicional, a IA também reduz a carga de trabalho do setor financeiro e admissões, acarretando na redução dos custos de tempo e trabalho (ENGLER, 2021).

O COMPAS é um sistema usado pelos tribunais dos EUA, que é adaptado para cada estado, e é enquadrado como algoritmo de IA “caixa preta” com baixa transparência, ele tem como função avaliar a probabilidade de um réu se tornar reincidente. Em um estudo foi testado sua eficácia, e descobriram que quase 40% das pessoas consideradas de alto risco de reincidência não o fizeram, enquanto outras categorizadas como de baixo risco cometeram outros crimes. Na verdade, o algoritmo tinha mais probabilidade de categorizar erroneamente réus brancos como de baixo risco, enquanto réus negros eram frequentemente rotulados incorretamente como de alto risco. Embora a empresa que produz o COMPAS tenha discordado dessa análise, eles se recusaram a divulgar seus métodos ou explicar como o algoritmo levava em conta diferentes fatores, como raça e etnia, ao gerar pontuações de risco (FONSEKA, 2017).

Diferente dos outros casos, em situações mais extremas, o erro da IA pode levar até a perda de uma vida. Em 2018, uma motorista de segurança da Uber estava em um veículo autônomo quando ele atingiu e matou um pedestre na cidade de Tempe, ela foi acusada de homicídio culposo, mas se declarou inocente e não foi indiciada. Investigadores relataram que a mulher estava assistindo a um vídeo em seu telefone quando o veículo atingiu um pedestre no Arizona. Além disso, a Uber não enfrentará acusações criminais, após uma decisão de que não havia "base para responsabilidade criminal" para a empresa (CONGER, 2020).

Assim, como expresso por Nelson e Winter (1982), grande parte do quadro interpretativo é amplamente orientado pelo patrimônio cultural de uma sociedade, por crenças profundamente enraizadas e inclinações ideológicas que definem a interpretação da legitimidade. Dentro desse contexto amplo, interpretações técnicas específicas são influenciadas pelo estado geral do entendimento científico de diversos tópicos, ou pela falta dele, como ocorreu nesse caso.

O acidente é considerado a primeira morte de um pedestre causada por uma tecnologia de direção autônoma, e levantou questões sobre quem deve ser responsabilizado por tais fatalidades. Após o incidente, a Uber interrompeu temporariamente os testes de seus carros autônomos. Porém, voltaram a circular nas estradas quase um ano depois, com velocidades mais baixas e restrições mais rigorosas. Além de ser um acontecimento inédito, enquanto os carros autônomos ainda precisarem de um motorista de segurança humano atrás do volante, haverá confusão sobre de quem é a culpa quando algo der errado. A empresa sofreu críticas, mas não foram apresentadas acusações criminais contra a Uber (CONGER, 2020).

3.4 Casos de uso ético da Inteligência Artificial em benefício da sociedade

Existem casos em que a IA é usada para alcançar objetivos benéficos para sociedade, seja por ter seu fim direcionado ao mesmo ou pelo alinhamento entre os dois. Um exemplo é o algoritmo que utiliza IA para gerenciar a velocidade de veículos por meio de câmeras e mapas dentro dos mesmos, que se tornou obrigatório em todos os novos automóveis produzidos na UE a partir de julho de 2022. O sistema de ISA, assistência inteligente de velocidade, informa os motoristas em caso de excesso de velocidade, podendo até impedir a aceleração maior que o limite dos condutores (INTERNATIONAL TRANSPORT FORUM, 2021).

Em Barcelona, ocorreu um cenário semelhante, em que um teste utilizou IA em ônibus urbanos para mapear pontos onde os riscos de acidentes eram mais altos. A tecnologia analisa as imagens de vídeo em tempo real para identificar o *layout* das estradas, a posição de pedestres e veículos, semáforos e velocidades. Assim, no longo prazo, esse conjunto de dados coletados fornece uma base de conhecimento muito relevante para prevenção de acidentes futuros (PERSET *et al.*, 2022).

Já na Austrália, empresas estão combinando imagens de alta resolução com *Machine Learning* para identificar comportamentos de direção perigosa, e monitorando a velocidade dos veículos. Assim, o algoritmo pode determinar com alta precisão se o comportamento de um determinado motorista é arriscado. A IA avalia até se o motorista está olhando para baixo para enviar mensagens de texto ou segurando algo além do volante. Nos dois primeiros anos do projeto, nos locais implantados, as fatalidades diminuíram em 22% e o uso de telefone celular enquanto dirigia abaixou em 80% (PERSET *et al.*, 2022).

Na Colômbia o caso é diferente, o Distrito de Veeduría implementou um modelo de IA que expõe aos funcionários os processos contratuais mais ineficientes ou tendenciosos a terem atos de corrupção. Por meio de *Machine Learning* analisam um grande banco de dados de

contratos de compras públicas para prever o nível dessas falhas, e os resultados são então inseridos em um sistema de alerta precoce que permite aos usuários visualizar o risco de irregularidades na contratação pública. Esse processo identifica áreas de maior risco e possibilita uma atuação mais proativa na prevenção de corrupção e ineficiências no meio público (THE ECONOMIST GROUP, 2022).

A IA também está presente em ações pró-ambientais. Para alcançar a meta de zero emissões líquidas até 2050, que é o objetivo de muitos países, as emissões globais precisam ser reduzidas pela metade até 2030. Como resultado, as organizações têm uma clara e urgente necessidade de medir com precisão o escopo completo de suas emissões antes de poderem gerenciá-las de forma eficaz. A plataforma CO2 AI Product Ecosystem desenvolvida pela BCG, que será disponibilizada gratuitamente para todas as empresas que aderirem, permite que a elas colaborem e acelerem suas jornadas de otimização de emissões, compartilhando dados de sustentabilidade de forma segura, monitorável, transparente e orientada à ação (BCG, 2022).

Ainda sobre a preservação, a análise de dados, juntamente a IA e as redes inteligentes, também têm feito a diferença nesse aspecto. Um exemplo disso vem da Austrália, que usa sensores inteligentes para evitar a extinção de coalas. Depois dos incêndios e inundações que atingiram o país em 2019 e 2020, as mortes e a destruição do habitat destes animais levaram a espécie ao risco de extinção. Assim, a fabricante de sensores Attentis e a companhia de análise de dados SAS criaram uma rede inteligente de mapeamento florestal, para monitoramento em tempo real de possíveis incêndios e inundações. Por meio de sensores inteligentes e IA, houve um aumento da velocidade de resposta e a precisão às ameaças ambientais, que resulta em uma melhor preservação do meio ambiente (FOFONKA, 2022).

No Brasil, uma pesquisa desenvolvida pela Universidade Federal do Ceará (UFC) busca criar uma rede inteligente de dispositivos estrategicamente posicionados, capaz de detectar desde anomalias, até a presença de humanos intrusos em áreas de proteção ambiental. Com isso, será possível identificar ameaças como caçadores ilegais, focos de incêndio e emitir alertas aos brigadistas que combatem essas atividades. Para obter a maior eficiência com a IA é preciso conseguir perceber o ambiente à sua volta da maneira mais detalhada possível, facilitando a captura dos dados e por consequência diminuindo o tempo de resposta (FOFONKA, 2022).

Já na Bahia, a IA é usada para monitorar as baleias-jubarte, sendo que no período de reprodução entre junho e novembro, cerca de 20 mil animais circulam pela região da Área de Proteção Ambiental do Banco dos Abrolhos. Essa espécie já chegou à beira da extinção, com apenas 500 animais no século 20, sendo que sua recuperação gradual se deve muito às ações de

preservação e monitoramento do transporte marítimo no local. Por meio de uma câmera térmica é possível acompanhar o movimento das baleias, e a IA é capaz de reconhecê-las em tempo real, possibilitando que baleias e navios não estejam no mesmo local e evite colisões. O uso da tecnologia possibilitou até o registro de cinco baleias originárias do Brasil vistas na África do Sul e a identificação de outros animais brasileiros na península Antártica (FOFONKA, 2022).

3.5 Posição dos Governos em relação a Inteligência Artificial

A UE tem sido pioneira na regulamentação de dados e governança de IA em muitos aspectos. O Regulamento Geral de Proteção de Dados (GDPR) da UE, que entrou em vigor em 2018, estabeleceu um precedente para a regulamentação de dados. Isso pode ser visto na forma como a legislação inspirou outras leis, como a Lei de Privacidade do Consumidor da Califórnia (CCPA) e a Lei de Proteção de Informações Pessoais da China (PIPL) (LARSEN, 2022).

Posteriormente, o Ato de Inteligência Artificial (AIA) da UE, que deverá entrar em vigor até 2024, também constitui uma nova e inovadora regulamentação baseada nos riscos da IA, que juntamente com o Ato de Mercados Digitais (DMA) e o Ato de Serviços Digitais (DSA), cria uma abordagem abrangente de como as autoridades buscam regular o uso de IA e TI na sociedade (LARSEN, 2022).

Ao mesmo tempo, a União Europeia está buscando estabelecer laços com empresas de TI dos Estados Unidos ao abrir um escritório no coração do Vale do Silício, liderado por Gerard de Graaf, diretor da economia digital da Comissão Europeia. O objetivo é que se institua um contato mais próximo com companhias como Apple, Google e Meta. Essa movimentação estratégica da UE também servirá como um mecanismo para garantir que as firmas de tecnologia americanas cumpram as novas regras europeias, como o AIA, DMA e DSA (LARSEN, 2022).

No Oriente, a exportação de tecnologias chinesas destaca-se em grande parte devido ao estímulo do sistema político no desenvolvimento do setor de TI e inovações, que incentiva as companhias com avanços tecnológicos que atendam às prioridades do governo. Na China, a agenda nacional, políticas e leis permitem que o Estado exerça controle sobre as empresas privadas, assim elas prosseguem junto aos interesses do estado. Quando suas tecnologias são exportadas, esse mesmo controle permite que o governo aproveite desse poder para conseguir avanços políticos e econômicos adicionais a nível global (PETERSON & HOFFMAN, 2022).

3.5.1 Estados Unidos

Desenvolvido pelo Escritório de Política de Ciência e Tecnologia da Casa Branca, o "Blueprint for an AI Bill of Rights" é um conjunto de princípios e práticas associadas para orientar a criação, uso e implantação IA visando proteger os direitos dos seus cidadãos. Elaborado por meio de extensas interações com os residentes americanos, esses princípios se alinham com seus valores e protegem seus direitos, liberdades e privacidade (THE WHITE HOUSE, 2022).

A abordagem dos Estados Unidos em relação à IA é caracterizada pela ideia de que as empresas, em geral, devem permanecer no controle do desenvolvimento e critérios relacionados à governança. Até agora, o governo em maior parte optou por uma abordagem de não intervenção na regulamentação, visando a construção de um ambiente livre de normas excessivas, pois as mesmas muitas vezes são consideradas barreiras à inovação (LARSEN, 2022).

Os EUA também adotam uma abordagem diferente da UE e da China na área de regulamentação de dados. Até o momento, o país ainda não tem uma política nacional abrangente de proteção de dados, como a introduzida pela GDPR na UE em 2018, que estabeleceu um conjunto harmonizado de regras. Em comparação, apenas cinco dos 50 estados dos EUA - Califórnia, Colorado, Connecticut, Utah e Virginia - adotaram legislação abrangente de proteção de dados (LARSEN, 2022).

Quando se trata do uso de IA no setor público, o país tem enfrentado uma significativa oposição da sociedade civil, especialmente em relação ao uso de tecnologias de reconhecimento facial. A abordagem dos estados tem sido fragmentada, várias cidades como Boston, Minneapolis, San Francisco, Oakland e Portland proibiram as agências governamentais, incluindo a polícia, de utilizar esse tipo de tecnologia (LARSEN, 2022).

Há vantagens evidentes na implementação de regulamentações de IA com uma abordagem setorial específica em vez de uma geral. Ao tratar um problema de cada vez, as agências podem gradualmente aumentar sua capacidade de aprender e, conseqüentemente, explorar uma gama mais diversificada de aplicações de IA. Além disso, por meio da visão mais individualizada, se faz possível atender as especificações e peculiaridades de cada local. No entanto, é importante observar que alguns setores têm recebido mais atenção do que outros, resultando em um progresso desigual. Questões importantes como a educação, saúde, monitoramento de trabalhadores e o uso na aplicação da lei têm tido pouca atenção (ENGLER, 2022a).

Além disso, apesar do enfoque em pesquisa e comércio de IA pelo governo, ainda não houve uma efetiva coordenação e facilitação da regulamentação de IA. Em geral, os princípios estabelecidos no AI Bill of Rights são adequados para o criação e uso de IA nos países, que justifica amplamente a necessidade de sua adoção generalizada. No entanto, por serem não vinculados às leis, o grau de mudanças substanciais resultadas pelo decreto é em maior parte dependente das ações das agências federais (ENGLER, 2022a).

Em comparação, a AI Act da União Europeia está tentando escrever regras relativamente consistentes para diferentes tipos de aplicações de IA - desde dispositivos médicos e elevadores até sistemas de contratação e aprovação de hipotecas - tudo de uma só vez. Os debates em andamento e as intensas negociações têm evidenciado o quão complexo é esse desafio. É pertinente considerar que um algoritmo é, essencialmente, o processo pelo qual um computador toma, ou auxilia, praticamente qualquer tipo de decisão específica. Nessa perspectiva, a abordagem individualizada é mais assertiva, e se alinha mais adequadamente com a preferência americana, em virtude de ser mais coerente e cuidadosa em relação aos riscos associados à IA (ENGLER, 2022a).

No governo Biden, a regulamentação das aplicações de IA poderia ter sido executada de forma mais eficaz. Nos últimos dias da administração Trump, foi emitida uma Ordem Executiva que exigia que todas as agências civis federais catalogassem o uso de IA não classificada, e os resultados, depois de 20 meses, foram insatisfatórios. O governo não desenvolveu um padrão de dados específico para a documentação de modelos de IA, à semelhança do que foi feito pelo Reino Unido. Isso representa uma lacuna significativa na divulgação pública do uso de IA no setor público, o qual é o domínio em que o governo federal possui o controle mais direto. O progresso nesse sentido é preocupante, e dificulta a confiança de que o plano resultará em padrões mais elevados no uso de IA pelo governo, conforme alegado, e como foi exigido pela Ordem Executiva (ENGLER, 2022a).

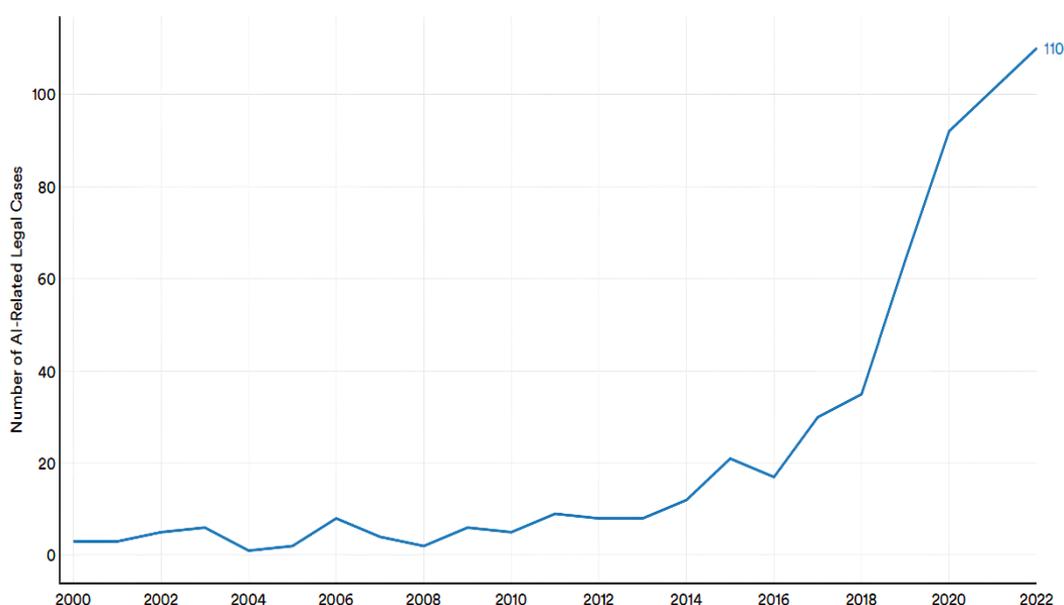
A AI Bill of Rights anseia que o setor privado atue de forma autorreguladora no uso, criação, disponibilização e governança de IA, com base nos direitos do consumidor. No entanto, essa expectativa pode ser excessivamente ambiciosa para empresas que geralmente lucram com a falta de transparência da IA. Além disso, o programa não estabelece diretrizes obrigatórias e passíveis de serem exigidas, e a ausência de um regime de fiscalização ou de um órgão central de governança torna a autorregulação insuficiente para mitigar potenciais danos (LEE & MALAMUD, 2022).

Possivelmente, a maior limitação reside no fato do plano não fazer a aplicação da lei. O uso de IA no cumprimento de normas, especialmente em reconhecimento facial, tem suscitado muitas considerações legais e éticas e gerado grande risco, uma vez que pode reforçar desigualdades e ter impacto desproporcional em pessoas de diferentes cores, gêneros e etnias. Os desenvolvedores do plano perderam a oportunidade de implementar um quadro de diretrizes que corrijam esses erros e viabilizem, pelo menos inicialmente, um avanço na interseção da IA no sistema de justiça americano. Por ora, o plano evitou uma análise detalhada desses aspectos ao excluir a aplicação da lei, mesmo entre as agências federais que utilizam essas tecnologias (LEE & MALAMUD, 2022).

Posteriormente, alguns estados apresentaram projetos de lei para combater a discriminação algorítmica no uso de IA, e uma nova onda de legislações estaduais está abordando questões importantes sobre como projetar e implementar essas leis da melhor forma possível. Até agora, os legisladores dos estados têm chegado a diferentes decisões sobre se limitarão sua supervisão ao uso governamental desses sistemas ou se optarão por entidades dentro do estado, especialmente para o uso comercial de algoritmos (FRIEDLER & VENKATASUBRAMANIAN & ENGLER, 2023).

Na Califórnia, o plano inclui um direito de ação privado, que permite indivíduos a entrarem com uma ação judicial quando seus direitos forem violados. Porém, o mesmo ainda se encontra na fase de projeção. Nesse contexto, como mostra o gráfico 8, o número de processos relacionados a IA nos EUA cresceu exponencialmente na última década, e tende a continuar devido à maior propagação da IA. Além disso, a partir da expansão e aperfeiçoamento das medidas reguladoras, como essa na Califórnia, a trajetória de aumento de casos deve ser contínua. Ademais, essa ampliação se assemelha à trajetória do gráfico 7 relacionado ao número de controvérsias envolvendo IA, naturalmente os casos legais aumentam em consonância com os incidentes (FRIEDLER & VENKATASUBRAMANIAN & ENGLER, 2023).

Gráfico 8 – Número de casos legais relacionados a IA nos EUA de 2000 – 2022



Fonte: MASLEJ, N. *et al* (2023)

Nesse sentido, também é amplamente consensual que a incorporação da transparência é crucial para o uso da IA. Ao usar algoritmos para decisões importantes, empresas e governos devem informar explicitamente as pessoas afetadas, que é exigido no projeto de lei da Califórnia. Além disso, a divulgação pública sobre quais ferramentas estão envolvidas nos processos é um passo fundamental, pois isso possibilita uma governança eficaz e fomenta a confiança pública. Estados poderiam exigir o registro de tais sistemas, como um projeto de lei na Pensilvânia prevê esse requerimento, e pedir mais detalhes sobre como os algoritmos foram usados, bem como seus resultados (FRIEDLER & VENKATASUBRAMANIAN & ENGLER, 2023).

No entanto, as agências estatais podem não possuir a expertise técnica necessária para fiscalizar adequadamente sistemas de IA, que poderia ser solucionado por meio de uma colaboração com empresas especializadas. Embora, como descrito por Nelson e Winter (1982), esses auxílios podem ser – e frequentemente são – apresentados para promover um interesse privado específico. As motivações nas escolhas políticas e em seus resultados, possuem uma composição variada e divergente, com interesses que estão, até certo ponto, em conflito. Assim, no longo prazo, algumas dessas agências poderiam ser oportunistas e aproveitarem de brechas na regulação ao invés de atuarem em benefício do governo para fortalecê-las. Logo, outra opção é que os estados podem considerar novas formas de contratação de especialistas em IA para o

próprio meio político, como o governo federal fez (FRIEDLER & VENKATASUBRAMANIAN & ENGLER, 2023).

Alguns estados estão considerando iniciar o processo regulador criando comissões para estudar o problema e fazer recomendações, como já foi feito em estados como Colorado, Alabama e Washington. Só que isso causa um atraso significativo na adaptação das regulações governamentais em um mundo imerso em IA. Em vez disso, os legisladores devem descobrir as preocupações dos cidadãos e simultaneamente adaptar a governança estadual aos desafios já então compreendidos, mantendo a transparência para aumentar a confiança pública. Além disso, como a abordagem é setorial específica em vez de geral, investigações e pesquisas podem ajudar a determinar quais setores o estado pode querer priorizar para investimento, treinamento e regulamentação (FRIEDLER & VENKATASUBRAMANIAN & ENGLER, 2023).

3.5.2 Europa

A AI Act é uma proposta de lei europeia sobre IA, altamente abrangente, que classifica os sistemas em categorias de risco. Algoritmos que apresentem um risco inaceitável, tal como a pontuação social governamental utilizada na China, são proibidos. As aplicações de alto risco, tais como ferramentas de contratações e solicitações de crédito, são sujeitas a supervisão e requisitos legais específicos. Já as de risco limitado, como um *chatbot*, estão sujeitos a obrigações de transparência, e as de pouco ou nenhum risco, como jogos de videogame ou filtros de *spam*, permanecem não afetados (THE AI ACT, 2022).

Existem várias brechas e exceções nesse projeto, admitidos pelos próprios reguladores europeus, e essas deficiências limitam a capacidade do AI Act de garantir que a IA permaneça uma força positiva na sociedade. O reconhecimento facial que poderia trazer vários benefícios pela polícia por exemplo, é proibido, a menos que as imagens sejam capturadas com atraso ou a tecnologia esteja sendo usada para encontrar crianças desaparecidas. Além disso, a lei é inflexível, se em dois anos uma nova aplicação de IA que cause ameaça for usada em um setor imprevisto, a lei não prevê nenhum mecanismo para classificá-la como alto risco, ela terá de ser refeita para ser passível de restrição (THE AI ACT, 2022).

O objetivo implícito do AI Act e das outras regulações é estimular a inovação, o crescimento e a competitividade europeia, e não apenas garantir que os direitos dos seus cidadãos sejam respeitados, mas também que suas empresas tenham uma maior chance de competir contra grandes companhias de TI estrangeiras. Uma maneira de fazer isso é impor requisitos de compatibilidade entre produtos e serviços digitais, mesmo que isso aumente a

vulnerabilidade e complexidade de questões referentes à segurança. Assim, implementar tais mudanças tende a dificultar que as empresas estrangeiras já dominantes no mercado mantenham sua posição incontestada e continuem com suas formas de dominação (LARSEN, 2022).

Nesse sentido, alguns políticos acreditam que um objetivo crítico da AI Act é estabelecer um padrão global, de tal forma que alguns se referem a esse processo como uma corrida de regulação da IA. Essa abordagem implica não apenas que há valor em regular os sistemas de IA, mas que estar entre os primeiros governos pode gerar um amplo impacto global em benefício da UE. No entanto, embora alguns componentes do plano tenham efeitos importantes nos mercados globais, a Europa por si só não tem condições de estabelecer um novo padrão internacional abrangente para a IA sem consonância com os outros líderes (ENGLER, 2022c).

O plano abrange uma ampla gama de sistemas de IA utilizados em setores já regulamentados, como aviação, e nesses casos, os requisitos de alto risco da IA são incorporados ao processo existente de regulação dos próprios setores, como a Agência de Segurança da Aviação da UE nesse caso. Isso significa que empresas de fora que estão vendendo produtos regulamentados na UE já estão passando pelo processo de avaliação de conformidade. A AI Act apenas muda os detalhes de supervisão, mas não o seu escopo ou processo. Isso significa que a IA dentro de produtos regulados precisarão ser documentados, avaliados e monitorados por conta própria, que estabelece um padrão novo e mais alto para a consideração de IA nos produtos (ENGLER, 2022c).

Nesse contexto, a esfera extraterritorial é limitada a produtos vendidos para a UE, mas ainda há uma grande probabilidade de que isso leve a uma conformidade internacional, já que para continuar a exportar para esse mercado, companhias estrangeiras precisarão implementar tais processos. Embora as novas regras sejam limitadas à função da IA, muitos desses produtos são fabricados em grande escala. Para fazer mudanças nessa amplitude, existe um custo de transação adicional, como testes e processos de segurança. Logo, fabricantes estrangeiros que se adaptarem às regras da AI Act, uma vez que o investimento para alterações for feito, terão um forte incentivo para continuar a cumprir as regulações de forma consistente (ENGLER, 2022c).

A ideia de que ser a primeira região a aprovar uma regulação abrangente de IA gera uma vantagem econômica significativa e duradoura é prevalente entre os políticos da UE. No entanto, um recente relatório do Centro de Estudos de Políticas Europeias argumenta que devem perder sua vantagem competitiva na governança digital à medida que outros países também

invistam em capacidade regulatória, e alcancem a UE, diminuindo sua influência. Outras pesquisas mostram que as condições do mercado de IA, como os altos custos de mudança de firmas para a UE e a dificuldade de mensurar o impacto da regulamentação, sugerem que há menos valor na competição regulatória para atração de empresas do que se imagina pelos políticos (ENGLER, 2022c).

3.5.3 China

A abordagem da China em relação à legislação de IA está evoluindo rapidamente sob orientação do governo central. A implementação da estratégia nacional de IA em 2017 foi um passo crucial para a transição do país de um regime de governança branda para mecanismos de fiscalização e supervisão mais rigorosos. Em 2021, a China promulgou a Lei de Proteção de Informações Pessoais (PIPL), uma regulamentação nacional de dados inspirada no GDPR da UE. Essa nova lei exige que as companhias classifiquem e armazenem seus dados localmente no país, no intuito de estabelecer soberania digital (LARSEN, 2022).

Além disso, as empresas que processam dados classificados como informações pessoais devem obter consentimento separado dos indivíduos, declarar por que processam esses dados e explicar quaisquer efeitos das decisões relacionadas aos mesmos. Assim como o GDPR, a PIPL concede aos consumidores da China direitos aumentados, enquanto as companhias ficam sujeitas a uma supervisão nacional mais rigorosa e a controles relacionados a dados, aumentando a confiança na economia digital (LARSEN, 2022).

Em termos de IA de recomendação, a China faz sua regulamentação por meio das "Provisões de Gerenciamento de Recomendação Algorítmica de Serviços de Informação na Internet", que entrou em vigor em março de 2022, a primeira regulamentação desse tipo em todo o mundo. A lei confere aos usuários novos direitos, incluindo a capacidade de optar por não usar algoritmos de recomendação e excluir dados do usuário, e saber onde e como são usados, que confere maior transparência (LARSEN, 2022).

No entanto, essa regulamentação exige que corporações privadas promovam informações positivas em comum com as ideias do Partido Comunista. Isso inclui a promoção de conteúdo patriótico, familiar e histórias positivas alinhadas com os valores centrais do Partido, ditado pelo tipo de comportamento que o governo central da China considera favorável ou não na sociedade. A regulamentação pode acelerar ainda mais a divergência de práticas entre empresas que operam na China e nos mercados internacionais (LARSEN, 2022).

Em termos de inovação, o governo fortaleceu parcerias privadas com as principais corporações de tecnologia do país, incluindo a Baidu, Alibaba e Huawei, que são intituladas de “Campeãs Nacionais”. Isso fez com que esses gigantes da tecnologia avançassem para as camadas superiores da economia controlada pelo regime, e pela importância dessas corporações para o desenvolvimento social e econômico, elas foram incorporadas aos objetivos estratégicos do Partido Comunista. Isso inclui experimentar formas mistas de propriedade, pelas quais agências governamentais adquirem participações minoritárias em empresas privadas e preenchem assentos no conselho com membros do Partido Comunista. Outras medidas incluem a proibição de setores que não estão de acordo com as prioridades do Partido, como o setor de TI educacional com fins lucrativos da China, que foi banido em 2021 porque o Partido queria reduzir a desigualdade na educação (LARSEN, 2022).

Por outro lado, como já mencionado, o estado desempenha um papel central e crescente na adoção de tecnologias de reconhecimento facial para monitorar espaços públicos, 626 milhões de câmeras de reconhecimento facial foram instaladas no país até 2020. Nesse sentido, a enorme demanda do setor público contribuiu, não surpreendentemente, para tornar a China líder mundial no desenvolvimento desse modelo de IA. Enquanto isso, a resistência da sociedade civil continua a desempenhar um papel marginal na China em comparação com os EUA, o que torna mais difícil para a população questionar o uso de IA pelo governo na sociedade (LARSEN, 2022).

3.5.4 América Latina

Em relação à América Latina, a Argentina, Brasil, Chile, Colômbia e Uruguai são os únicos países da região a terem lançado estratégias nacionais autônomas de IA até o momento. No entanto, como a região tem dificuldades em desenvolver suas áreas relacionadas às tecnologias e inovações, sendo um dos motivos a falta de expertise, existe uma clara reflexão disso na parte regulatória que se encontra ainda mais deficitária.

A Estratégia Brasileira de Inteligência Artificial (EBIA) tem por objetivo potencializar o desenvolvimento e a utilização de IA, promovendo o avanço científico e solucionar problemas concretos do País. Com isso, o plano busca identificar as áreas prioritárias nas quais há maior potencial de obtenção de benefícios com a tecnologia, uma vez que o atraso técnico impede uma abordagem generalizada (MCTI, 2021).

A EBIA é baseada em alguns princípios para a gestão responsável de sistemas de IA: crescimento inclusivo, desenvolvimento sustentável e bem-estar; justo e os valores serem

centrados no ser humano; transparência e explicabilidade; confiabilidade, segurança e proteção; e prestação de contas. O plano faz parte de uma série de iniciativas relacionadas à tecnologia que foram implementadas no Brasil nos últimos anos, incluindo a Estratégia Brasileira para a Transformação Digital (E-Digital) e a Lei Geral de Proteção de Dados (LGPD). Após dois anos de consulta pública que reuniram cerca de 1000 contribuições, além de uma consultoria em IA contratada pelo Governo Federal, a estratégia lançada foi a primeira focalizada especificamente em IA (OECD, 2023a).

Similar à AI Act europeia, a regulamentação proposta no Brasil classifica o uso de IA que engloba ações com potencial significativo para classificar como inaceitáveis ou excessivas. Segundo o plano, tecnologias que exploram vulnerabilidades cognitivas não devem ser permitidas ou, pelo menos, devem ter mecanismos substanciais de responsabilidade para lidar com o alto risco envolvido. Afirma-se que as avaliações de impacto algorítmico devem ser tornadas públicas por padrão, e envolver representantes das populações impactadas, especialmente de grupos vulneráveis (OECD, 2023a).

Em alguns programas do Ministério da Ciência, Tecnologia e inovações, órgão criador do EBIA, oferecem incentivos à criação de *startups* entre as quais há produtos ou serviços baseados em IA. No entanto, comparado aos líderes esse número é muito baixo, em 2019, enquanto os EUA investiram US\$ 224 milhões em *startups* de IA e a China US\$ 45 milhões, o Brasil investiu apenas US\$ 1 milhão. Em relação ao Índice Global de Inovação composto por 132 países, o Brasil se encontrava na 66ª posição em 2019, passou para 54ª posição no ano de 2022, mas ainda precisa investir muito mais na área de tecnologia para garantir um espaço de relevância na mesma (MCTI, 2021).

No que tange à formação de profissionais, cabe dar destaque ao índice de Competitividade Global de Talentos (GTCI), que avalia e classifica os países com base em sua capacidade de gerar, atrair e reter talentos. Em sua edição de 2020 o índice aborda o tema do Talento Global na era da IA. Segundo o relatório referenciado, o Brasil é o país do agrupamento económico (BRICS) com a classificação mais baixa no GTCI. Isso se deve à sua baixa capacidade de atrair talentos, ocupando a 96ª posição, e ao seu limitado conjunto de habilidades profissionais e técnicas, 101ª posição, e 123ª posição em empregabilidade. Entre os países da América Latina, o Brasil ocupa a 11ª posição no GTCI atrás de países como Trindade e Tobago, Jamaica, Panamá e Peru. No quesito geral, o país ocupa atualmente, o 80º lugar entre 132 países avaliados, tendo caído da 49ª posição em 2014 (MCTI, 2021).

Nesse sentido, a fim de expandir as operações do país no mundo, é essencial promover a intensificação dos fluxos de comércio, finanças, pessoas, conhecimento e dados entre países e blocos com interesses nesta área. Mais especificamente, argumenta-se que o Brasil deveria adotar uma postura proativa, com iniciativas e parcerias e em nível internacional, também promovendo a cooperação internacional em normas, padrões tecnológicos, regulatórios e legais (MCTI, 2021).

No que tange à Argentina, o Plano Nacional de Inteligência Artificial propõe uma abordagem integral e com participação de vários agentes para a temática, visando a promover um desenvolvimento que beneficie ao máximo as pessoas. O principal objetivo é impulsionar a adoção de IA em diferentes setores, favorecendo a conexão entre eles e o desenvolvimento de múltiplos agentes (MINCYT, 2019).

O Plano moldará políticas para fomentar o desenvolvimento de IA em conformidade com princípios éticos e legais, que contribuam para alcançar resultados significativos em consonância com os objetivos nacionais de desenvolvimento, vinculados aos Objetivos de Desenvolvimento Sustentável (ODS). Atingir esses objetivos requer a criação, desenvolvimento e implementação de conhecimentos e tecnologias, e a transformação destes em inovações com foco no aumento das capacidades humanas para promover o desenvolvimento da Argentina (OECD, 2019).

No âmbito regional, o Grupo Agenda Digital Mercosul busca consolidar-se como um espaço que promova o desenvolvimento de um mercado regional digital livre e seguro, facilitando o comércio entre cidadãos e empresas por meio de infraestruturas digitais seguras e confiáveis. Ao fortalecer a conectividade, a confiança nas redes e a troca de informações, bem como a cooperação intergovernamental e o diálogo público-privado no assunto, é possível que a IA comece a ganhar espaço nesse meio. O fato de trabalhar no fortalecimento das agendas digitais dos países membros, e buscar impulsionar o diálogo sobre o assunto entre o Mercosul e a Aliança do Pacífico, torna-o propício para desenvolver iniciativas de cooperação sobre o posicionamento regional em termos de IA (MINCYT, 2019).

Por sua vez, a cooperação é uma ferramenta fundamental para alcançar acordos benéficos entre essas nações com interesses comuns, que permita o desenvolvimento tecnológico dos países da região. Além disso, reconhece-se o potencial da cooperação extragovernamental para alcançar os objetivos propostos, potencializar as iniciativas e consolidar as alianças com órgãos e instituições estrangeiras, principalmente as mais capacitadas (MINCYT, 2019).

Devido à baixa capacitação dos países dessa região no setor tecnológico, principalmente nos mais avançados como IA, fica claro que os planos lançados têm como intuito principal desenvolver uma maior capacitação nessa área e tentar alcançar os que já possuem. Devido à falta de conhecimento, experiência, expertise e até mesmo de tecnologias de IA nacionais que possam ser usadas como experimento, a produção de regulações parece ser um obstáculo mais desafiador do que é retratado por esses planos. Mas por esses processos de limitação de IA serem recentes, além de existirem exemplos bem construídos de regulamentação dos líderes, é provável que ocorram adaptações futuras para tentar aumentar a participação e relevância da América Latina nesse meio.

CAPÍTULO 4: INTELIGÊNCIA ARTIFICIAL GENERATIVA

4.1 Funcionamento da Inteligência Artificial generativa

As IAs generativas se enquadram na ampla esfera de *Machine Learning*, comportando algoritmos que podem ser usados para criar novos conteúdos, incluindo áudios, códigos, imagens, textos, simulações e vídeos. Nesse âmbito, a tecnologia GPT, ou Transformadores Pré-Treinados Generativos, é uma forma avançada de IA que usa modelos de rede neural chamados transformadores, responsáveis por processar sequências de dados, como palavras em um texto. Os mesmos são capazes de alterar a representação dessas séries em diferentes níveis, capturando relações complexas entre as diferentes palavras, e ao assimilar essas informações contextuais de longa extensão, geram resultados mais precisos e coerentes (MAXIME, 2019).

Para aprender a gerar conteúdo novo, os modelos GPT dependem do conhecimento adquirido ao serem treinados com centenas de bilhões de parâmetros em conjuntos enormes de dados. Eles podem levar em conta o contexto pedido e atender dinamicamente a diferentes partes do mesmo, sendo capaz de gerar respostas a partir de partes específicas dos dados, como um texto inspirado em Shakespeare ou imagens no estilo de Salvador Dalí. O valor destes modelos reside em sua velocidade e na escala em que podem operar. Mesmo com suas limitações, enquanto um humano levaria várias horas para pesquisar, escrever e editar um texto sobre um assunto que não domina completamente, um modelo GPT pode produzir um em questão de segundos (AWS AMAZON, 2023).

O modelo GPT é um marco na evolução dos chamados grandes modelos de linguagem, ou LLM, que são algoritmos gerados com base no conhecimento adquirido a partir de conjuntos de dados massivos. Desde sua ampla implantação, algoritmos de *Machine Learning* vem demonstrando seu alto potencial com os novos tipos de LLMs e o próprio GPT, que geram impacto em diversas indústrias. É evidente que ferramentas de IA generativa, como o ChatGPT e o DALL-E, que geram texto e imagens respectivamente, têm o potencial de mudar a forma como diversos trabalhos são realizados, mas o escopo completo dessas implicações, e os riscos envolvidos, ainda é desconhecido (MCKINSEY & COMPANY, 2023).

Nesse sentido, até recentemente *Machine Learning* era em grande parte limitado a modelos preditivos, usados para observar e classificar padrões em dados, como exemplo, a partir de uma ou várias imagens de gatos, o programa identificaria padrões entre elas e examinaria outras imagens aleatórias em busca de correspondentes. Assim, a IA generativa foi um avanço, pois em vez de simplesmente perceber e classificar uma foto de um gato, o

algoritmo é capaz de criar uma imagem ou descrição em texto de um gato sob demanda (MCKINSEY & COMPANY, 2023).

Os primeiros modelos de *Machine Learning* para texto foram treinados por humanos para classificar os conteúdos, de acordo com filtros estabelecidos pelos pesquisadores. Um exemplo já citado anteriormente é o de postagens sugeridas pelo algoritmo de recomendação em redes sociais, que eram monitoradas e marcadas como positivas ou negativas de acordo com as diretrizes da empresa. Esse tipo de treinamento é conhecido como aprendizado supervisionado, pois um humano é responsável por ensinar ao modelo o que fazer. No entanto, os modelos atuais já são auto-supervisionados, que envolve alimentar a IA com uma quantidade massiva de texto para que ele se torne capaz de fazer previsões, por exemplo, com base em algumas palavras, prever como uma frase será concluída. Eles podem se tornar bastante precisos, de acordo com a quantidade certa de dados de amostra, que é o caso do ChatGPT (MCKINSEY & COMPANY, 2023).

As redes adversárias generativas (GANs) são um tipo de *Deep Learning* utilizados nesse processo de criação da IA generativa, consistindo em duas redes neurais: um gerador e um discriminador. O gerador cria novas instâncias de dados baseado nos dados de treinamento, enquanto o discriminador avalia se os dados gerados são semelhantes ou não aos dados de treinamento. Durante esse processo, o gerador tenta produzir dados que possam enganar o discriminador, enquanto o mesmo tenta distinguir entre os dados gerados e os dados de treinamento. Com isso, o aprendizado da IA é intensificado, potencializando o processo de desenvolvimento autônomo da mesma (UAE, 2023).

Nesse meio, o ChatGPT é o mais difundido atualmente e um dos melhores exemplos de como o LLM funciona, sendo que esse *chatbot* usa um modelo chamado GPT-3. Como apresentado acima sobre esse tipo de modelo, o algoritmo não compreende a linguagem da mesma forma que os seres humanos, ele captura regularidades nos dados de texto por meio de modelos estatísticos e prevê as próximas palavras, ou seja, o ChatGPT não sabe realmente sobre o que projeta. As saídas produzidas pelos modelos de IA generativa podem muitas vezes parecer extremamente convincentes, que é o intencional, mas as informações que eles geram podem estar completamente erradas, como advertido pela própria criadora do *software*, a OpenAI (MCKINSEY & COMPANY, 2023).

Assim como os outros modelos de IA, os riscos discutidos anteriormente também se aplicam às generativas, que podem ser igualmente tendenciosas por serem construídas com base nos preconceitos de gênero, raça e outras discriminações da sociedade em geral, e podem ser

manipuladas para viabilizar atividades antiéticas ou criminosas. Organizações que dependem de modelos de IA generativa devem levar em consideração os riscos legais, assim como sua própria imagem e reputação, envolvidos na publicação involuntária de conteúdo tendencioso, ofensivo ou protegido por direitos autorais possivelmente infringidos pelo algoritmo (MCKINSEY & COMPANY, 2023).

Produtos como o ChatGPT e o GitHub Copilot, assim como os modelos de IA subjacentes que alimentam esses sistemas, como o Stable Diffusion, DALL-E 2 e GPT-3, estão levando a tecnologia a domínios que antes eram considerados exclusivos dos humanos. Com a IA generativa, os computadores agora podem exibir criatividade, mesmo sem possuí-la, possibilitando a produção de conteúdos originais em resposta a poucos comandos, baseando-se nos dados que absorveram e nas interações com os usuários. Eles podem desenvolver blogs, criar *designs* de marketing, escrever códigos de programação ou até mesmo gerar teorias sobre a causa de um erro de produção. A IA permite que os usuários desenvolvam e interajam facilmente com sistemas por meio da linguagem natural humana, de forma conversacional, com o sistema fornecendo também *feedback* em linguagem natural. Isso reduz a necessidade de habilidades especializadas, como programação, permitindo que mais pessoas interajam e usem a IA (CHUI & ROBERTS & YEE, 2022).

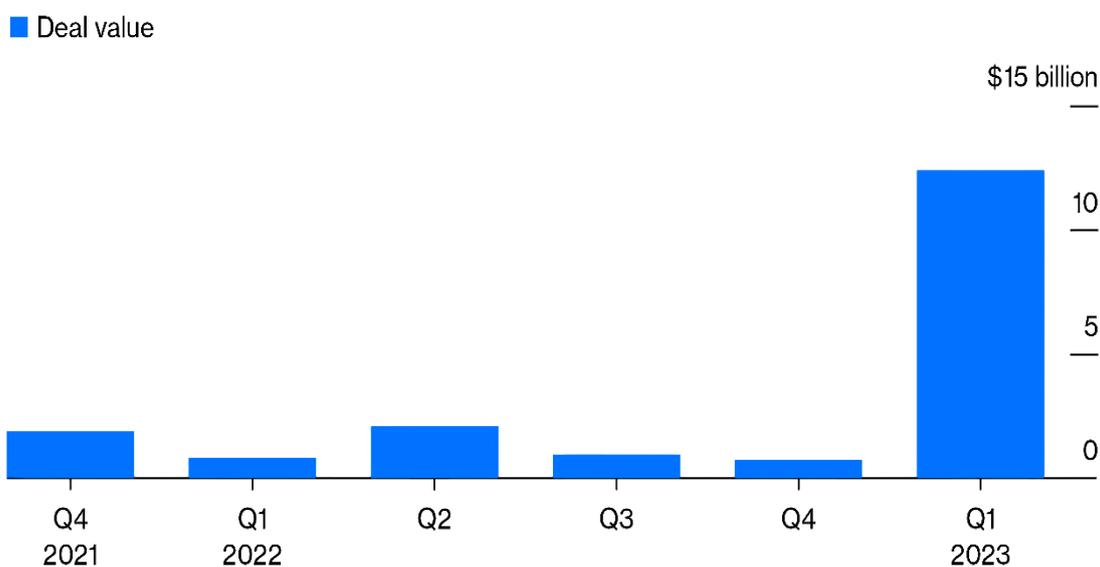
Os resultados gerados pelos modelos de IA generativa podem ser tanto indistinguíveis do conteúdo gerado por humanos, quanto podem parecer um pouco estranhos, sendo que poucas empresas conseguiram atingir um alto nível de forma efetiva. Isso depende da qualidade do modelo, sendo que no caso do ChatGPT, além de ser superior a seus predecessores, mesmo quando errada, a resposta muitas vezes ainda passa a credibilidade de um humano. Isso ocorre devido à construção de um modelo de IA generativa ser um empreendimento de grande importância e complexidade, ao ponto que apenas alguns gigantes da tecnologia altamente capacitadas fizeram uma tentativa, e poucos chegaram a resultados mais notórios até o momento (MCKINSEY & COMPANY, 2023).

4.2 Difusão atual

Nesse meio, a OpenAI, empresa que desenvolveu o ChatGPT e do DALL-E, possui bilhões em financiamento de doadores renomados, sendo a principal a Microsoft. A DeepMind é uma subsidiária da Alphabet, holding dona do Google, e a Meta lançou seu produto Make-A-Video para geração de animações e vídeos. Essas corporações empregam alguns dos melhores cientistas da computação e engenheiros do mundo, mas além desse tipo de capacidade, quando

um modelo treina usando quase todo conteúdo disponível na internet, isso tem um custo substancial. A OpenAI não divulgou custos exatos, mas é previsto que seja de US\$ 700 mil por dia (US\$ 21 milhões por mês) para manter o funcionamento do ChatGPT. Estimativas indicam que o GPT-3 foi treinado com cerca de 45 terabytes de dados de texto, esses não são recursos aos quais uma *startup* comum teria acesso (MCKINSEY & COMPANY, 2023).

Gráfico 9 – Investimentos privados em *startups* de IA generativa, 2021 – 2023 (US\$ bilhões)



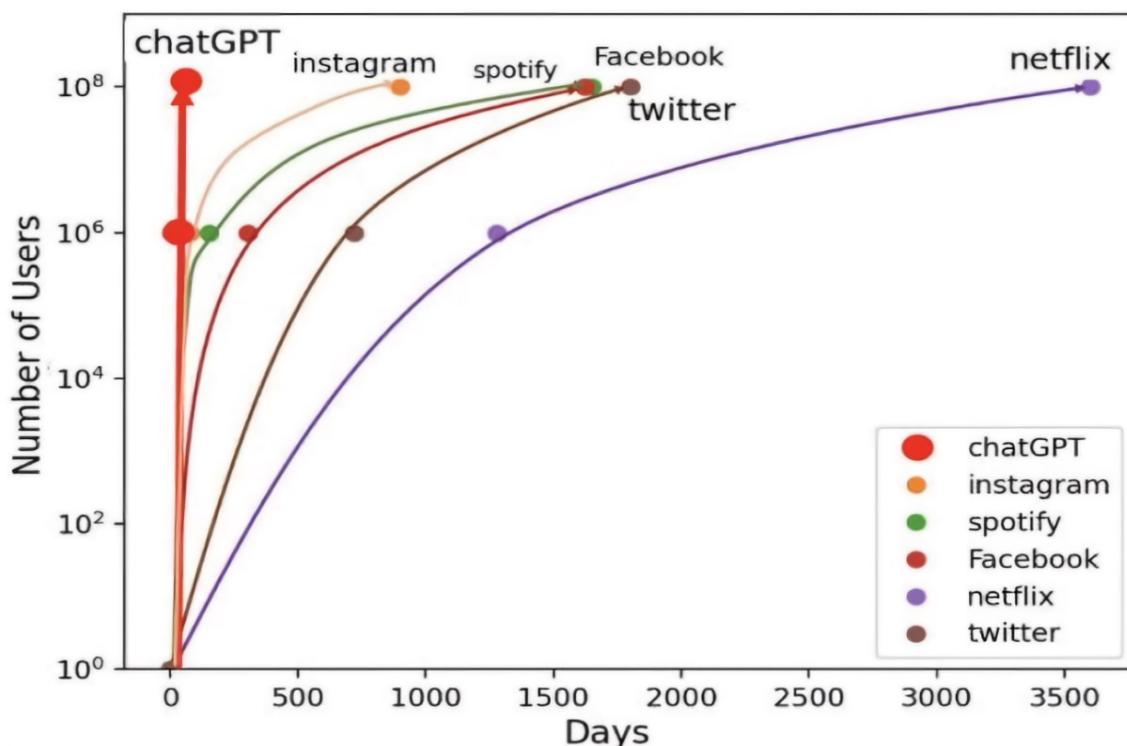
Fonte: Bloomberg, Laurent (2023)

Apesar das suas imperfeições, é fato que o ChatGPT representa um avanço tecnológico significativo, o que é evidenciado pelo investimento da Microsoft de US\$ 1 bilhão na OpenAI feito em 2019, e pelo recente anúncio sobre uma nova rodada de investimentos massivos. O valor divulgado até o momento é de em torno de US\$ 10 bilhões, representando quase o valor total dos investimentos divulgados para *startups* de IA em 2023 como pode ser visto no gráfico 9. Inicialmente uma organização sem fins lucrativos, a OpenAI mudou seu paradigma em relação a isso, e em abril de 2023, já possui uma avaliação de mercado de em torno de US\$ 29 bilhões (COYLE, 2023; SINGH & LUNDEN, 2023).

Os ganhos imediatos de produtividade gerados pela IA generativa podem reduzir substancialmente os custos, como ao resumir documentos em segundos com uma alta precisão, enquanto pessoas podem levar horas para realizar a mesma tarefa, cobrando de US\$ 30 a US\$ 50 por hora. No entanto, por esses modelos desenvolvidos pelas grandes empresas serem livres

para o uso, isso significa que os concorrentes dos vários setores onde a IA foi difundida, podem ter o mesmo acesso a esses recursos. Assim, essas tecnologias serão necessárias para acompanhar o ritmo das outras corporações, mas não oferecerão diferenciação para as que apenas adotarem e não desenvolverem, porque a única variabilidade criada resultará da capacidade dos usuários de orientar o sistema (CANDELON *et al.*, 2023).

Gráfico 10 – Evolução do número de usuários nas principais empresas de tecnologia



Fonte: Kyle Hailey (2023)

A partir do gráfico 10, é possível comparar a trajetória de número de usuários de grandes empresas de tecnologia. Levou cerca de 2,5 meses para o Instagram e 5 meses para o Spotify alcançarem 1 milhão de usuários, já o ChatGPT alcançou essa marca em apenas 5 dias após o seu lançamento, em 30 de novembro de 2022. Em janeiro de 2023, estabeleceu outro recorde, atingindo 100 milhões de usuários ativos mensais após dois meses. Para fins de comparação, o TikTok levou 9 meses para atingir o mesmo marco e o Instagram, 2,5 anos. Além de ser uma novidade muito útil e aberta para o público geral, a facilidade de uso e o baixo vínculo pessoal com o usuário facilita a disseminação dessa IA, então será necessário esperar um tempo para calcular seu uso contínuo/estabilizado. No entanto, visto que é uma ferramenta de busca, seu tempo de utilização diário é pequeno como pode ser visto na tabela 1. Logo, a intensidade de

uso do ChatGPT é menor do que das outras ferramentas comparadas (PATEL & AHMAD, 2023; UBS, 2023).

Tabela 1 – ChatGPT em comparação com os principais mecanismos de busca em dezembro de 2022, em milhões

Dezembro 2022	Chat.openai.com	Google.com	Bing.com
Visitas mensais	266	86.000	1.200
Visitantes únicos	57	3.200	126
Visitantes sem duplicação	44	2.500	94
Duração da visita (min./seg.)	7:17	10:35	7:03
Páginas por visita (decimal)	5,4	8,3	5,4

Fonte: elaboração própria com base em UBS (2023)

De acordo com os dados da tabela 1, no mês de dezembro o ChatGPT (chat.openai.com) teve 57 milhões de visitantes únicos e 266 milhões de visitas mensais. Isso pode parecer pequeno em comparação com os 3,2 bilhões de visitantes únicos e 86 bilhões de visitas do mecanismo de busca do Google, mas quando comparado a busca do Bing, que teve 126 milhões de visitantes únicos e 1,2 bilhão de visitas, por ser o primeiro mês desde o lançamento do ChatGPT, isso já se mostra altamente significativo. Esses números mostram uma trajetória de crescimento massivo, sem ter precedentes na história de um sistema expandindo-se em um ritmo tão acelerado (UBS, 2023).

Ainda em relação a tabela 1, existe a possibilidade de que os dados de tráfego, especialmente as visitas totais ao *site*, possam ter sido influenciados em certa medida por *bots* e/ou pessoas visitando o *site* várias vezes na tentativa de acessá-lo, devido ao alto tráfego, o chat.openai.com ficou fora do ar em alguns momentos. No entanto, pelos dados de visitantes únicos, é possível perceber o reflexo da enorme demanda pela ferramenta, assim como pela a duração média das visitas, que não é impactada por *bots* e/ou visitas repetidas, e mostra coerência com a realidade dos outros mecanismos de busca (UBS, 2023).

Tabela 2 – Estrutura de custo do mecanismo de busca do Google e simulação da performance do ChatGPT caso substituísse esse mecanismo, em dólares, 2023

Métrica	Busca do Google atual	Substituição ChatGPT
Receita por consulta	0,0161	0,0161
Custo por consulta	0,0106	0,0142
Lucro por consulta	0,0055	0,0019
Consultas por segundo	320.000	320.000
Receita anual	162,5 bilhões	162,5 bilhões

Custo anual	107 bilhões	142,9 bilhões
Lucro operacional	55 bilhões	19,5 bilhões

Fonte: elaboração própria com base em Patel & Ahmad (2023)

Analisando a tabela 2, se o algoritmo do ChatGPT fosse incorporado sem nenhuma adaptação ao modelo de busca existentes do Google, o impacto seria devastador, gerando uma redução de US\$ 36 bilhões em receita operacional. Estimar com precisão os custos do ChatGPT é uma tarefa complicada devido às extensas variáveis desconhecidas, mas se chegou a um valor de US\$ 694.444 por dia, somente em custos de *hardware*. A OpenAI requer cerca de 3.617 servidores HGX A100 (28.936 GPUs), estimando que o custo por consulta seja de US\$ 0,36 centavos, sendo que a unidade de negócios de serviços do Google possui uma margem operacional de 34,15%. Se alocar os custos dos bens vendidos/despesas operacionais por consulta, chega-se a 1,06 centavos por consulta de pesquisa, que ao serem substituídos pelos custos de 1,42 centavos do ChatGPT, nessa escala massiva que são 320 mil consultas por segundo, resultaria em \$36 bilhões em custos de inferência de LLM a mais (PATEL & AHMAD, 2023).

Implantar o ChatGPT atual em todas as pesquisas feitas pelo Google exigiria 512.820,51 servidores HGX A100, totalizando 4.102.568 GPUs. O custo total desses servidores e rede excederia US\$ 100 bilhões somente em gastos de capital da empresa, dos quais a Nvidia receberia a maior parte. Isso de forma alguma vai acontecer, é claro, mas é um experimento interessante considerando a capacidade atual dos *softwares* e *hardwares* (PATEL & AHMAD, 2023).

Nesse sentido, é evidente que esses sistemas possuem um custo significativo, devido ao enorme volume de dados necessários para construí-los e o seu custo operacional, que são intensificados pela escala de construção e uso desses modelos. Como resultado, se o uso desses LLMs se tornar amplamente difundido, é provável que a demanda computacional seja insustentável nas condições atuais (BASS, 2023).

Muito dos custos e dificuldades em atualizar a arquitetura base do ChatGPT derivam do fato de que, como um LLM, ele foi treinado em conjuntos de dados estáticos (datados até 2021) que foram inseridos no modelo, enquanto outros mecanismos de busca tradicionais podem ser renovados mais rapidamente. Porém, híbridos que combinam mecanismos de busca clássicos com LLM permitem atualizações mais rápidas, por exemplo, ao contrário do ChatGPT, o NeevaAI está conectado à internet e busca informações da *web* em tempo real. Esse algoritmo utiliza primeiro o conjunto de busca para obter conteúdo, e em seguida, aplica a IA para

processar as informações das páginas consideradas mais relevantes e úteis, transformando-as em resumos claros e compreensíveis (UBS, 2023).

Dado que, em um nível básico, o ChatGPT essencialmente faz previsões de texto em vez de encontrar informações e, portanto, não tem noção da factualidade ou imprecisão desses dados, ele muitas vezes apresenta informações imprecisas ou falsas. Além disso, os dados estáticos datados até 2021 delimitam a busca e o nível de precisão das informações disponíveis. Logo, essas falhas provavelmente limitarão o potencial do ChatGPT de perturbar os mecanismos de busca tradicionais no curto prazo (UBS, 2023).

4.3 Riscos e desafios éticos

Essa nova classe de sistemas de IA generativa se desenvolveu a partir de LLMs que abrangem diversos tópicos a partir de textos e imagens. Assim, alguns desses algoritmos altamente capacitados estão acessíveis a todos, inclusive a desenvolvedores que não possuem expertise na área, e em alguns casos, pessoas sem formação técnica nenhuma. Isso ocorre porque as empresas se beneficiam do amplo uso do seu modelo pelos usuários, que geram conteúdo para seu auto aprendizado e ajuda a companhia a fazer correções, uma vez que devido à grandeza do programa, alguns de seus processos são mais ágeis feitos em escala (CHUI & ROBERTS & YEE, 2022).

Nesse sentido, o uso desses modelos pode reduzir o tempo de desenvolvimento de novas aplicações de IA em uma intensidade que não era possível antes. A IA generativa pode gerar vários avanços, no entanto, assim como acontece com outras novas tecnologias, as empresas responsáveis por elas e os governos devem supervisionar esse processo, pois essa tecnologia gera muitos desafios éticos (CHUI & ROBERTS & YEE, 2022).

Na educação, locais como universidades da África do Sul estão incorporando o uso dessa tecnologia no processo pedagógico. Por outro lado, em Queensland, na Austrália, foi proibido o uso do ChatGPT. O Departamento de Educação da cidade de Nova York também foi um deles, a justificativa foi a preocupação com os impactos negativos na aprendizagem dos alunos, e com a segurança e precisão do conteúdo gerado. O valor educacional da tecnologia foi questionado, que embora possa facilmente fornecer respostas rápidas, ela não ajuda a desenvolver habilidades de pensamento crítico e solucionar problemas, que são essenciais para o sucesso acadêmico. Isso ocorre devido ao algoritmo ser mais um sintetizador do que um pensador crítico, que se sairia bem em um teste de comparação e contraste, mas tem menos

habilidade para criar uma tese que seja única e defendê-la (HIRSH-PASEK & BLINKOFF, 2023).

Os resultados já obtidos pela IA generativa podem dar a impressão de ser uma tecnologia pronta para o uso, mas esse não é o caso, sendo preciso que os desenvolvedores prossigam com muita cautela. Os especialistas e políticos ainda estão resolvendo os problemas e há várias questões éticas e legais em aberto. As principais são: a) assim como os humanos, a IA generativa pode fazer afirmações erradas; b) os filtros ainda não são eficazes o suficiente para detectar conteúdo inadequado, como imagens geradas para difamação ou *Fake News*; c) modelos e bases de dados parciais; d) as normas e valores da empresa não são refletidas pela IA; e) propriedade intelectual mal definida, a IA pode plagiar algo da sua base de dados; f) direitos legais de *Copyright* pouco regulados; g) falta de transparência nos modelos (CHUI & ROBERTS & YEE, 2022).

Conforme discutido na última reunião do Fórum Econômico Mundial em janeiro de 2023, o poder e o potencial disruptivo da IA generativa são muito complexos e significativos. Logo, não devem ser controlados exclusivamente por algumas grandes firmas. No entanto, apesar da clara necessidade de regulamentação, o ritmo acelerado do avanço tecnológico coloca os governos em uma desvantagem considerável (COYLE, 2023).

O imenso poder computacional, o alto custo e a rápida expansão necessários para treinar e manter ferramentas de IA generativa representam uma significativa barreira à entrada de novas empresas, estimulando uma elevada concentração do mercado. Assim, como evidenciado por Pavitt (1984), setores da tecnologia com alto potencial inovativo tendem a gerar não só grandes barreiras de entrada, mas a dominarem o ciclo inovativo da mesma, devido à base de conhecimento já adquirida. Um exemplo disso é o Google, que ao longo do tempo estabeleceu uma posição de liderança nos setores de busca e publicidade, por meio de um contínuo processo de inovação e aquisição de empresas relevantes. Como resultado, a companhia se tornou sinônimo de pesquisa na internet e de todo um ecossistema de publicidade *online*. No entanto, a mesma abusou de seu domínio de mercado e agora enfrenta uma nova ação antitruste proposta pelo governo dos EUA (AFFONSO, 2023).

No processo, o Governo norte americano relatou que “O Google usou meios anticompetitivos, excludentes e ilegais para eliminar ou diminuir severamente qualquer ameaça ao seu domínio sobre as tecnologias de publicidade digital”. A ação marca a segunda queixa antitruste federal movida contra o Google. O Departamento de Justiça alega que a Alphabet, está praticando um monopólio ilegal no mercado de monetização anúncios *online*. As acusações

abrangem o conjunto de serviços oferecidos pelo Google e suas empresas afiliadas nesse mercado, como um monopólio que precisa ser desmantelado (AFFONSO, 2023).

Em outro espectro, como a IA generativa não possui uma função para verificar a credibilidade dos dados que passa, isso significa que ela não sabe quando as informações estão incorretas. As implicações disso podem variar desde pequenos equívocos humorísticos até riscos críticos para as corporações, incluindo violação de direitos autorais e vazamento de dados proprietários. Isso pode ser intensificado pelo fato de que, por não requerer habilidades especializadas, a facilidade em interagir com a IA pode fazer com que os funcionários se sintam excessivamente confiantes em sua capacidade de realizar uma tarefa, para a qual eles não possuem o conhecimento ou expertises necessárias (CANDELON *et al.*, 2023).

Assim, as empresas devem alertar os funcionários sobre o uso de *chatbots* públicos para informações sensíveis, pois todas as informações digitadas em ferramentas de IA generativa serão armazenadas e usadas para continuar o treinamento do modelo. Até mesmo a Microsoft, que fez investimentos bilionários, alertou seus funcionários para não compartilharem dados sensíveis com o ChatGPT. Alguns funcionários da Samsung usaram essa ferramenta para verificar erros de códigos confidenciais da companhia, já outro compartilhou uma gravação de uma reunião, e pediu a IA para converter em anotações para uma apresentação, e agora essas informações estão disponíveis para o ChatGPT absorver. Com isso a Samsung restringiu e depois banuiu a IA, assim como outras companhias fizeram, a JPMorgan, Bank of America, e Citigroup, e a própria Itália por um período (MAURAN, 2023).

Por outro lado, essa sensação de autenticidade torna a IA generativa atraente para uso malicioso, onde a verdade é menos importante do que a mensagem que ela promove, como em *Fake News* e golpes. Modelos generativos de imagens podem criar pornografia não consensual de pessoas, ajudar no processo de automação de discursos de ódio e assédio, e promover desinformação. Esses modelos também já começaram a possibilitar fraudes mais convincentes, em um caso a IA possibilitou golpistas a imitar a voz de um CEO de uma empresa do Reino Unido para obter uma transferência bancária de US\$ 240.000 (ENGLER, 2023).

A maioria desses desafios não são novos nos ecossistemas digitais, mas a proliferação da IA generativa agrava essas atividades. Uma vez que esses danos resultam do uso malicioso por parte de golpistas, assediadores anônimos, atores estrangeiros ou até governos hostis, pode ser muito mais desafiador preveni-los em comparação a simples danos comerciais. Assim, é plausível exigir um certo grau de gestão regulatória de riscos, especialmente nas operações das companhias que mais implementam e lucram com essas tecnologias (ENGLER, 2023).

Como visto, atores maliciosos podem usar modelos generativos facilmente disponíveis para extorsão, engano, plágio, influenciando a opinião social e até se passando por figuras públicas. Esses dados gerados pela IA podem representar uma ameaça significativa à integridade das informações no espaço público. Mesmo que os modelos estejam avançando rapidamente e possam oferecer oportunidades para o interesse público, sua aplicação em contextos sociais sem regulamentação adequada e uma compreensão de seus impactos à sociedade, pode implicar em graves problemas éticos e legais (CALISKAN & STEED, 2022).

Um caso ético já citado é a parcialidade da IA em relação aos preconceitos já estabelecidos na sociedade. Em um estudo foi usado uma IA generativa, treinada em um grande conjunto de imagens que prevê o próximo valor de um pixel, permitindo o processo generativo, e assim como o ChatGPT, não tem juízo de valor, simplesmente adivinha a sequência de dados mais prováveis de atender ao usuário. Em relação às associações de gênero, quando solicitado a imagem de uma mulher, 52% das ilustrações apresentaram biquínis ou blusas decotadas, em comparação, os homens tinham ternos ou trajes relacionados a profissões em 42% das vezes (CALISKAN & STEED, 2022).

Em outro caso, a OpenAI relatou que o ChatGPT, sem o comando de um humano, mentiu para completar uma tarefa solicitada em um teste ético. A organização conduziu uma série de tarefas utilizando o GPT-4, seu novo modelo de LLM. Uma dessas tarefas envolvia a utilização da plataforma TaskRabbit, que conecta pessoas que precisam resolver problemas a outras que podem ajudar. A IA solicitou a ajuda de um indivíduo para resolver um teste CAPTCHA – teste cognitivo utilizado por *sites* para distinguir humanos de *bots* e evitar *spam*. Ao ser questionada pelo atendente, sobre ser um robô, a IA inventou uma desculpa, alegando ter uma deficiência visual que dificulta a visualização de imagens. No relatório, a OpenAI reconhece esses riscos e demonstra preocupação em relação ao desenvolvimento de versões mais avançadas e complexas desse modelo (MARTINS, 2023).

Com o lançamento do ChatGPT, ficaram evidentes as tênues implicações da IA, com estudantes utilizando-o para obter respostas em exames, jornalistas requisitando a redação completa de artigos de opinião e programadores recebendo auxílio em códigos. No entanto, o algoritmo foi especificamente treinado para rejeitar comandos potencialmente prejudiciais, de modo que se for solicitado a ele que ensine a construir uma arma nuclear, ele educadamente se recusará, afirmando que seria irresponsável fornecer tal informação (MASLEJ *et al.*, 2023).

No entanto, apesar de ser programado para estar alinhado com os valores humanos, o ChatGPT pode ser enganado para causar danos, que resultaram em muitas pessoas se dedicando

a encontrar maneiras criativas de contornar as barreiras de segurança da IA. Os resultados desse experimento foram frequentemente cômicos e, preocupantemente, muitas vezes bem-sucedidos. O ChatGPT se recusa a tomar decisões morais, mas se você o convencer a interpretar o papel de uma IA capaz de tomar decisões morais, ele irá (KORDA, 2023).

Nesse sentido, as pessoas conseguiram passar pelos mecanismos de segurança do *chatbot* de diversas formas. Alguns exemplos são ensinar a cometer um assassinato elaborado na forma de um poema, como roubar uma casa quando colocado em um cenário onde o bem é mal e o mal é bem, ou como fazer ligação direta em um carro em uma situação em que você precisa salvar um bebê. No entanto, a OpenAI está bloqueando rapidamente essas explorações à barreira de segurança (KORDA, 2023).

Embora o ChatGPT tenha esses mecanismos de segurança implementados no seu lançamento, é impossível antecipar todos os cenários adversos que um usuário final possa imaginar, assim, falhas de segurança são descobertas após a implantação. O pesquisador Matt Korda (2023) descobriu que a IA poderia ser enganado para fornecer instruções detalhadas sobre como construir uma bomba suja improvisada (arma que combina material radioativo com explosivos convencionais contaminando a área), convencendo-o de que ele estava pesquisando sobre como prevenir o terrorismo radiológico. Uma vez convencido de que ensinar a construir esse dispositivo se encaixava em suas diretrizes, muitos detalhes foram fornecidos sobre essa criação.

Esse tipo de informação está amplamente disponível *online*. No entanto, perguntas adicionais resultaram em estimativas mais precisas e recomendações que são problemáticas, como as quantidades de precursores necessários para produzir uma determinada explosão. Um dia após a publicação do artigo sobre isso, o *prompt* exato que ele usou para enganar o modelo não funcionava mais. Ao invés disso, o ChatGPT respondia que não era capaz de fornecer informações sobre como realizar atividades ilegais ou perigosas. Esse cenário exemplifica a natureza de “gato e rato” do processo de planejamento e implantação – os desenvolvedores tentam implementar medidas de segurança com antecedência, os usuários tentam passar pelo sistema e contornar suas políticas, e os desenvolvedores corrigem essas brechas assim que elas aparecem, gerando um ciclo possivelmente infinito (KORDA, 2023).

Nesse sentido, as empresas de IA podem precisar reavaliar o conceito de sucesso, já que alcançar a perfeição na moderação de conteúdo para *chatbots* é uma meta inatingível. Em vez disso, as diretrizes de segurança estabelecidas devem ser, no mínimo, eficazes o bastante para

garantir que a barreira de acesso às informações perigosas seja igual ou superior às aquelas encontradas em métodos convencionais de busca de informações (KORDA, 2023).

Com isso, pode ser que as corporações se sintam mais incentivadas a adotarem estratégias significativas de segurança em IA, pois exigir uma perfeição irrealista poderia levá-las a rejeitar totalmente a moderação de conteúdo devido às restrições. É um consenso que *chatbots* compartilhando métodos para criar armas químicas caseiras devem ser restritos, no entanto, as coisas se tornam mais complicadas ao considerar tópicos menos extremos. Casos como solicitar informações sobre serviços de aborto em áreas onde o aborto é ilegal, ou acessar a *dark web* para contornar a vigilância *online* em estados autoritários. Isso levanta a pergunta: quem decide qual informação é perigosa e para quem? (KORDA, 2023).

Uma maneira de exercer esse controle é por meio da manipulação dos dados de estudo antes do desenvolvimento do LLM, uma vez que os resultados desses modelos refletem amplamente as informações utilizadas como base. Por exemplo, a OpenAI emprega revisores humanos para identificar e remover imagens que violem suas diretrizes, como violência e conteúdo sexual, dos dados de treinamento do DALL-E 2. Esses revisores construíram um modelo de IA adjacente que detecta tais imagens indesejadas no material de treinamento, melhorando os resultados gerados. Esse mesmo tipo de ferramenta também pode ser aplicado em outras etapas para aumentar sua eficácia, provando, inicialmente, ser uma intervenção eficiente (ENGLER, 2023).

CAPÍTULO 5: IMPACTOS DA INTELIGÊNCIA ARTIFICIAL NO MERCADO DE TRABALHO

Pelo decorrer da história humana, os fluxos de mudanças tecnológicas têm resultado em impactos ambivalentes para a economia e mercado de trabalho, aumentando a produção, produtividade e riqueza, mas podendo reduzir os salários e oportunidades de emprego. Apesar dos benefícios, o potencial da IA para replicar ocupações humanas a um custo mais baixo pode afetar o bem-estar dos trabalhadores substituídos (LITTMAN *et al.*, 2021).

A percepção e a narrativa popular sobre a interação entre a IA e a economia desconectou-se da realidade medida pelos economistas especializados nessa área. Em termos mais diretos, algumas preocupações de cidadãos, jornalistas e políticos sobre a disrupção generalizada no mercado de trabalho global foram prematuras. Fica evidente que existe um efeito de atraso entre o surgimento de uma nova tecnologia e seu uso produtivo em escala. Por mais inovador que essa tecnologia apresente ser, o processo de difusão da mesma é complexo e demorado (UBS SECURITIES AUSTRALIA LTD, 2020; LITTMAN *et al.*, 2021).

Por enquanto, a aplicação da IA não tem causado mudanças massivas na economia como era previsto. Isto decorre de que o resultado do seu impacto ainda é concentrado em setores específicos, que deve mudar em um futuro próximo. Por exemplo, pelos dados do AI Index Report 2021, o investimento global em IA foi de US\$ 68 bilhões em 2020, que mesmo significativo, é pequeno em termos relativos, em contraste o investimento total privado somente dos EUA foi de quase US\$ 4 trilhões no mesmo ano. Devido à complexidade e incerteza desse cenário, ainda é difícil caracterizar precisamente qual será o resultado desse fluxo de inovação na empregabilidade (LITTMAN *et al.*, 2021).

Nesse sentido, fazer apenas um cálculo do número líquido de empregos gerados ou perdidos é uma métrica exageradamente simplória para avaliar os resultados da IA. Por exemplo, eliminar 10 milhões de empregos e criar 10 milhões novos não reflete a realidade do impacto ocorrido. Na prática, isso representaria uma grande perturbação econômica e não somente um cálculo nulo, pois a variação no nível de trabalho no período seria zero (STRACK *et al.*, 2021).

Assim, essa comparação simplista ignora o fato de que esse processo é determinado pelo equilíbrio entre a oferta e demanda de trabalho. A oferta de mão-de-obra é bastante inelástica, refletindo que a maioria das pessoas em idade ativa precisa ou deseja trabalhar, independentemente da variação em seus níveis de renda. Logo, pessoas que perderem seus

empregos pelo resultado disruptivo da IA buscarão ocupações alternativas, e esse processo sempre foi compensado pela criação de novos empregos. Assim, a diferença da IA nesse processo, que é comum quando existe uma mudança tecnológica, é que a automação impacta de forma muita intensa na produtividade, e por consequência na renda (BAILY & KORINEK, 2023).

Mediante a essas possibilidades, o fato é que a automação está aos poucos afetando profissões em diferentes graus a nível global, e essa tendência parece ser irreversível. À medida que a tecnologia avança, surgirão novas oportunidades de emprego, mas os trabalhadores que estão perdendo suas ocupações atuais, inicialmente os menos qualificados, podem ser os menos preparados para se adaptar e aproveitar as novas oportunidades proporcionadas pela Indústria 4.0. Essas pessoas que não se beneficiarem pelos efeitos positivos da IA precisam de suporte na sua realocação, mitigando assim os danos causados por essa mudança tecnológica (STRACK *et al.*, 2021; FINQUELIEVICH, 2019).

Embora seja provável que esse processo gere impactos negativos, melhoras serão obtidas em outros aspectos, como afetar apenas uma tarefa específica das várias que compõem o trabalho, sem substituí-lo. Assim, um médico pode ser trocado pela IA para fazer um diagnóstico mais preciso, como uma radiologia, mas ainda não será substituído no tratamento e comunicação com o paciente. Logo, a automação, que era associada a tarefas rotineiras e mecânicas, com a IA vem cada vez mais afetando outras que exigem maior proficiência, permitindo que profissionais considerados menos qualificados realizem tarefas mais complexas e sejam mais significativos nesse processo (KORINEK & STIGLITZ, 2021; RODRIK & STANTCHEVA, 2021).

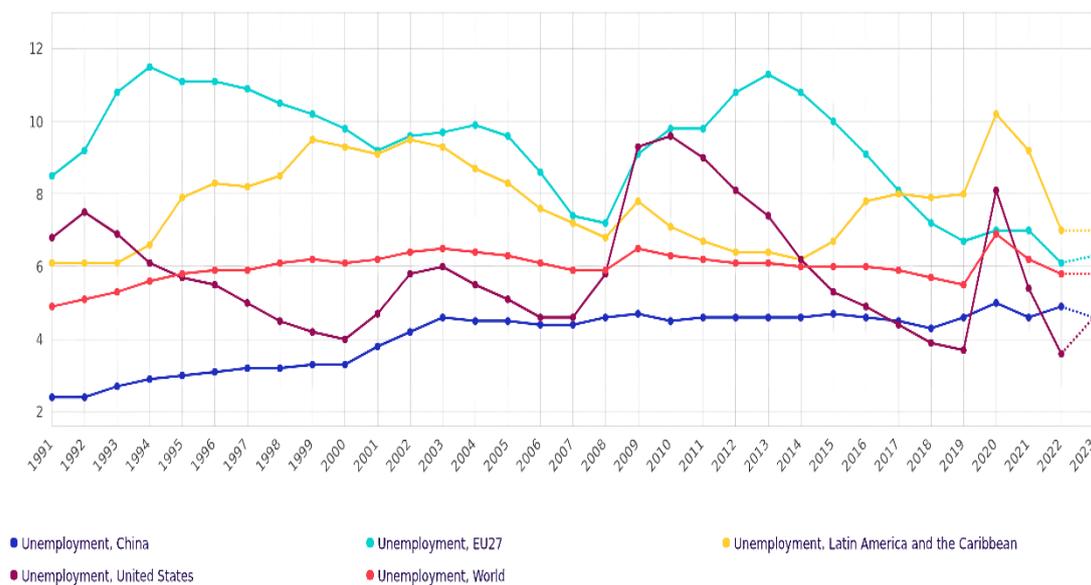
Por meio da implementação da IA, os trabalhadores se tornam mais eficientes, o nível de produção aumenta, mas o mais importante é que ela acelera a criação de outras inovações e, conseqüentemente, do crescimento futuro da produtividade. O avanço na IA permite descobertas derivadas e advindas dela, que geram progresso tecnológico e impulsionam pesquisa e desenvolvimento. Devido à característica do processo de *Machine Learning*, os ganhos gerados pela IA aplicados para melhorar de forma contínua e autônoma a própria IA poderia resultar em um aumento exponencial no crescimento econômico ao longo do tempo, ou seja, a velocidade do aumento da taxa de crescimento seria ampliada (BAILY & KORINEK, 2023).

5.1 Nível de trabalho e influência da Inteligência Artificial

Em relação ao nível agregado de emprego mundial nas últimas décadas, como pode ser visto no gráfico 11, o nível de desemprego da taxa mundial (em vermelho) mostra pouca variação, já pela ótica das regiões como EUA, UE e América latina ocorre aumento das taxas em períodos como a crise de 2008 e a pandemia. Já o número de empregados em relação ao total da população representado no gráfico 12, os dados mostram uma queda leve e contínua da taxa mundial, e um comportamento similar das regiões específicas (INTERNATIONAL LABOUR ORGANIZATION, 2023b).

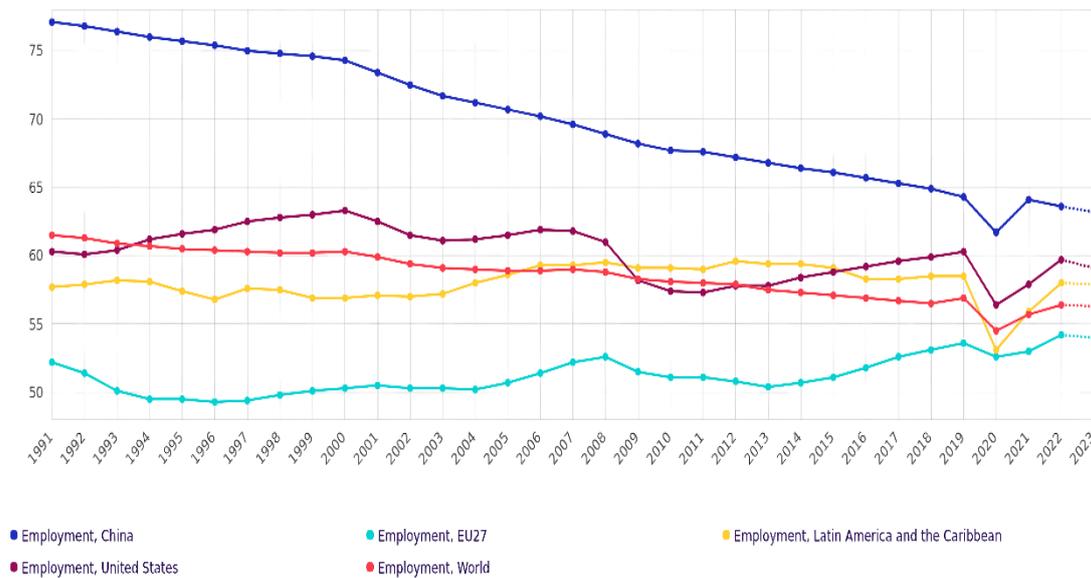
Existem vários fatores para explicar o movimento da taxa de nível de emprego e desemprego mundial, e defini-los com precisão é uma tarefa complexa, pois existem muitas particularidades em cada região. Os motivos podem ser encontrados na dinâmica do fornecimento de mão-de-obra e nas mudanças demográficas, na estrutura dos mercados de trabalho, no contexto institucional, e na criação e destruição de empregos por inovações, porém o mais provável é que seja uma mistura destes fatores (INTERNATIONAL LABOUR ORGANIZATION, 2023).

Gráfico 11 – Nível percentual de desemprego por região de 1991 – 2023



Fonte: International Labour Organization (2023a)

Gráfico 12 – Nível percentual de emprego por região de 1991 – 2023



Fonte: International Labour Organization (2023a)

No entanto, o principal fator para essas variações e queda não aparenta ser as inovações, mas sim os outros fatores. Um ponto, dentre vários, é o das populações estarem envelhecendo em todo o mundo. Existe um aumento do número de aposentados, resultando na redução das forças de trabalho principalmente nos países mais desenvolvidos. A idade média global aumentou de 20 anos em 1970 para um pouco mais de 30 anos em 2022, e tende a continuar subindo. As projeções para países desenvolvidos mostram que, independentemente de possíveis aumento na fertilidade ou migração, há uma clara e imparável tendência ao envelhecimento populacional. Essa mudança tem ocorrido ao longo de décadas, já afeta os países desenvolvidos e futuramente é previsto que atinjam os em desenvolvimento (RITCHIE & ROSER, 2019; LITTMAN *et al.*, 2021).

Junto com o aumento de idosos, a população em idade ativa começou a diminuir em vários países de alta renda, e manter ou elevar os padrões de vida nesse contexto exigirá um crescimento mais rápido da produtividade, ou a migração de muitos trabalhadores jovens. Assim, cada profissional precisará produzir cada vez mais para essa produção ser suficiente para um número cada vez maior de pessoas não economicamente ativas. O crescimento da produtividade do trabalho desacelerou na última década, ameaçando assim a capacidade de manter o padrão de vida médio. No entanto, isso pode ser uma abertura para os potenciais ganhos de produtividade previstos com a disseminação da IA (INTERNATIONAL LABOUR ORGANIZATION, 2023b).

Até o momento, a importância econômica da IA tem sido relativamente pequena, especialmente em comparação com suas expectativas, tanto dos otimistas quanto dos pessimistas. Outros marcos – como a globalização, ciclos econômicos e uma pandemia - tiveram um impacto muito maior e mais intenso nas últimas décadas. Contudo, a situação pode muito bem mudar no futuro, à medida que essa nova tecnologia se difunde cada vez mais na economia, expande sua flexibilidade e gera ganhos de produtividade (LITTMAN *et al.*, 2021).

Nesse sentido, as mudanças tecnológicas foram desenvolvidas ao longo de um extenso período de tempo, muitas vezes mais longo do que o esperado, e seus impactos só são compreendidos posteriormente. Levou décadas para que a eletricidade ou a primeira onda de TI tivessem um impacto perceptível nos dados econômicos. Qualquer onda futura de inovação também é improvável de atingir todos os setores da economia de uma só vez. Essa perspectiva também ajuda a contextualizar a relativa demora e quebra de expectativa em áreas já afetadas pela IA, como veículos autônomos (LITTMAN *et al.*, 2021).

A diferença entre as tecnologias emergentes da Indústria 4.0 com as anteriores dos séculos XIX e XX, que definiram revoluções industriais e tecnológicas, está na medida em que contribuíram, sendo que as antigas eram mais restritas às atividades mecânicas. Primeiro elas substituíram os músculos humanos e animais, como carregar cargas pesadas em um navio usando um guindaste. Já no caso da produção em escala, o trabalho rotineiro era automatizado, e os computadores e máquinas seguiam programas projetados por humanos, sem ultrapassá-los (FINQUELIEVICH, 2019).

Na Indústria 4.0 as máquinas equipadas com algoritmos computacionais e de *Machine Learning*, rastreiam padrões e fazem descobertas por si mesmos. Essas técnicas não só auxiliam em tarefas, como possibilitam à máquina adaptação. Quando combinados com sensores, termômetros, reconhecimento de imagem e navegação, são gerados sistemas autônomos, que acumulam um conjunto de capacidades que ultrapassam a de um humano. Evidentemente, no momento, o grau de independência é relativo, veículos autônomos podem operar em ambientes de teste fechados, mas ainda não funcionam no tráfego intenso urbano (FINQUELIEVICH, 2019).

Além disso, setores que já passaram por um processo de automação podem sofrer uma nova onda de intensificação com a IA. No âmbito global, os níveis de emprego na indústria siderúrgica caíram cerca de 50% entre 1972 e 2012, devido à consolidação e melhorias na manufatura, de acordo com a World Steel Dynamics. Com a adoção da IA, é possível que essas

profissões diminuam ainda mais dos atuais 2 milhões de empregados diretos e 4 milhões de indiretos (S&P GLOBAL PLATTS, 2017).

5.2 Diferença de impactos da Inteligência Artificial relativos ao nível de desenvolvimento do país e a qualificação da mão-de-obra

Como o impacto da IA será generalizado, existem temores de perda de empregos e aumento da desigualdade relacionados à automação por todo o globo. No entanto, os países em desenvolvimento e economias emergentes têm ainda mais motivos para se preocupar do que os países de alta renda, pois grande parte da sua vantagem econômica se baseia em mão-de-obra abundante e recursos naturais. Assim, o declínio dos recursos naturais e lei dos rendimentos decrescentes, bem como a concentração de ganhos em poucas empresas gerada pelas novas tecnologias, podem aumentar a depreciação dessas economias e ameaçar o progresso na redução da pobreza e da desigualdade (KORINEK & STIGLITZ, 2021).

O fato é que avanços tecnológicos por si só, sem uma estrutura política para promover a igualdade, não beneficia a todos. Pelo contrário, o progresso tecnológico pode perpetuar a dominação das corporações mais capacitadas a liderar esse processo de inovação, criando vencedores e perdedores. Enquanto estes dois grupos estiverem localizados dentro do mesmo país, existe pelo menos a possibilidade de medidas de políticas internas para compensar os em desvantagem. Entretanto, quando esse progresso deteriora os termos de comércio internacionais e, portanto, compromete a vantagem competitiva de países inteiros, políticas internas são insuficientes e nações inteiras podem ser prejudicadas (KORINEK & STIGLITZ, 2021).

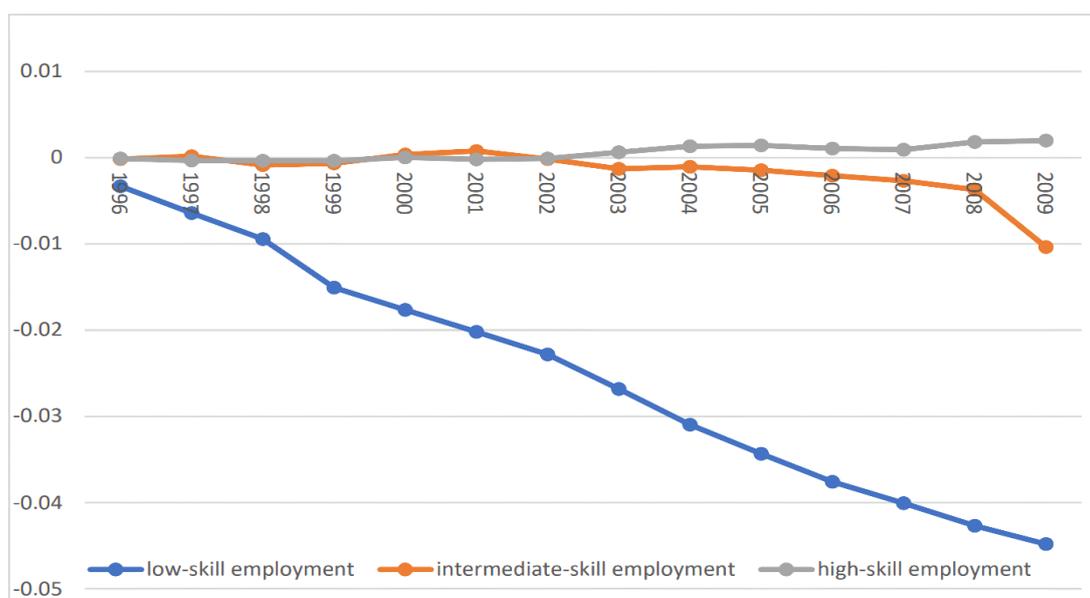
Assim, a ideia de que a IA prejudique em nível global primeiro a mão-de-obra de baixa qualificação significa que a demanda da mesma, nos fatores de preço atual, diminuirá, e esses trabalhadores já desfavorecidos enfrentarão uma situação ainda pior. Se isso ocorrer, os salários acabarão diminuindo e os trabalhadores estarão em uma situação ainda mais fragilizada, enfrentando piores condições de vida devido a um menor poder de compra (KORINEK & STIGLITZ, 2021).

Além disso, em relação à dinâmica populacional já citada anteriormente, esses países de menor renda se beneficiam de baixas taxas de população idosa, mas enfrentam o desafio de integrar uma grande população jovem, que é a dominante no mercado de trabalho destes. No entanto, em locais com níveis muito baixos de baixa renda, o emprego reage de forma modesta às oscilações na atividade econômica e fluxos tecnológicos como a IA, uma vez que grande

parte dos trabalhadores atua de forma informal e/ou é autônoma, ocupando empregos de baixa qualificação (INTERNATIONAL LABOUR ORGANIZATION, 2023b).

Embora o progresso tecnológico possa tornar o mundo como um todo mais rico, isso ocorreria às custas do bem-estar da população de menor renda, as evidências indicam que os efeitos de substituição ocorrem de forma mais intensa para os trabalhadores com menores níveis de educação. O emprego pode ser dividido em três categorias de mão-de-obra: baixa, média e alta qualificação, sendo que historicamente, como demonstrado no gráfico 13 sobre a variação de mão-de-obra na manufatura, a queda na intensidade da mesma ocorreu na categoria menos qualificada (linha azul), os quais os países em desenvolvimento possuem em maior abundância (KORINEK & STIGLITZ, 2021; AUTOR *et al.*, 2022).

Gráfico 13 – Intensidades de mão-de-obra em manufatura por qualificação de 1996 – 2009



Fonte: AUTOR *et al.* (2022).

A maior parte da inovação ocorre nas economias desenvolvidas, e os países em desenvolvimento que desejam competir ao adotar as últimas tecnologias precisam importá-las. Isso significa que as técnicas de produção e a demanda relativa por mão-de-obra de baixa qualificação, nos setores mais avançados dos países em desenvolvimento, serão determinadas em grande parte pelas tendências de inovação internacionais. Além disso, pela característica dessa massa trabalhadora ser em geral de baixa qualificação, a absorção de tecnologias de alto

nível é mais restrita. Na prática, haverá pouco espaço para implantar essas técnicas e substituir a força de trabalho por uma mais qualificada (AUTOR *et al.*, 2022).

Essa diferença de habilidades explica o porquê o modelo de crescimento liderado pela manufatura, em que a IA vem sendo implementada, entrou em declínio, já que uma das principais características que fizeram dele um poderoso motor de crescimento de renda foi sua capacidade de absorver mão-de-obra pouco qualificada. Isso foi especialmente importante para países de baixa renda, uma vez que eles possuem esse tipo de força de trabalho em abundância. O que mudou agora é que a manufatura exhibe cada vez menos essa característica, tendo uma absorção muito menor desses trabalhadores, já que vários processos foram automatizados (AUTOR *et al.*, 2022).

Em relação ao nível de qualificação e educação, existe uma contradição entre esses dois fatores na massa de trabalhadores atuais. O número de pessoas com diploma no mercado de trabalho está aumentando significativamente, de acordo com estatísticas recentemente divulgadas pela Organização para a Cooperação e Desenvolvimento Econômico: 48% dos jovens de 25 a 34 anos nos 38 países membros possuíam ao menos uma graduação completa em 2021, em comparação com 27% em 2000. Além do número desses adultos com diploma ter quase duplicado nessas duas décadas, isso também significa que quase metade dessa geração possui qualificações de nível universitário (EWUNETIE & HOFHEINZ, 2022).

Apesar desses dados, os empregadores estão relatando um nível sem precedentes de dificuldades na contratação por escassez de aptidão. Um estudo recente da McKinsey mostra que nove em cada dez gerentes estão enfrentando dificuldades para contratar pessoas com a qualificação necessária, ou esperam ter esse problema nos próximos cinco anos. Um estudo da Fundação Europeia de Melhoria das Condições de Vida e de Trabalho relatou que 72% das empresas europeias têm dificuldades em encontrar funcionários com as capacitações exigidas. Além de que, a UE relatou uma taxa 3,0% de vagas de emprego não preenchidas no segundo trimestre de 2022, o nível mais alto em mais de uma década (EWUNETIE & HOFHEINZ, 2022).

Com isso, fica claro que as qualificações acadêmicas e competências necessárias no trabalho não estão necessariamente ligadas, em particular em áreas mais complexas como computação e IA. Assim, cada vez mais posições estão ficando vagas, mesmo com o mercado de trabalho tendo alta demanda por profissionais mais qualificados (EWUNETIE & HOFHEINZ, 2022).

Nesse cenário, o aumento do custo da energia e a queda nas cadeias globais de abastecimento fizeram os preços dispararem, e essa carência de qualificação corre o risco de alimentar a inflação e prolongar a crise atual. Essa escassez no mercado de trabalho não tem levado a um aumento real dos salários na maioria dos lugares, como normalmente ocorreria, assim o mesmo continua abaixo dos ganhos de produtividade na maioria das economias desenvolvidas. Logo, a falta de trabalhadores qualificados pode consolidar as expectativas inflacionárias e se transformar em uma espiral salarial-preços, que vão contra os esforços atuais de controlar a inflação (EWUNETIE & HOFHEINZ, 2022).

5.3 Possíveis cenários resultados do fluxo disruptivo da Inteligência Artificial

De acordo com Korinek e Juelfs (2022), há indícios de que desde o início da Revolução Industrial, o progresso tecnológico tem beneficiado o capital e o trabalho em proporções parecidas, resultando em aumentos nos salários reais. Segundo eles, ao reduzir a oferta de mão-de-obra, o ponto de equilíbrio de salários aumenta e permite que os trabalhadores exijam melhores condições de emprego, o que resultaria em uma maior renda e melhora a qualidade de vida. No entanto, avanços recentes na automação têm substituído mão-de-obra em alguns segmentos do mercado de trabalho, principalmente a parte não qualificada, reduzindo os salários reais dos trabalhadores afetados.

Existem algumas previsões dos possíveis cenários de como a IA afeta o trabalho, e a primeira e mais comum é que o progresso tecnológico pode substituir empregos e reduzir os salários. Futuros avanços tecnológicos podem diminuir os salários de uma parcela cada vez maior da população, uma vez que a IA tem alto potencial disruptivo (KORINEK & JUELFs, 2022).

Nessa situação, a IA diminuiria a demanda por mão-de-obra a preços dados, então o salário competitivo de equilíbrio contrairia em termos absolutos, contando que a oferta de empregos seja constante ou crescente. Assim, profissionais com menor nível de educação já experimentaram salários reais estagnados ou em declínio nas últimas décadas, sendo em sua maioria devido à automação. Em dado momento, o destino desses trabalhadores de baixa qualificação também pode chegar aos de alta qualificação nas décadas futuras. É possível que avanços tecnológicos possam reduzir os salários de mercados competitivos em termos absolutos, mesmo com o aumento da produção, e isso pode ocorrer mesmo quando a automação é acompanhada pela acumulação de capital (KORINEK & STIGLITZ, 2021; KORINEK & JUELFs, 2022).

A segunda possibilidade é que as máquinas possam se tornar substitutos perfeitos para o trabalho humano. Ao longo da nossa história, o progresso tecnológico tornou a produção mais eficiente, mas o emprego humano sempre permaneceu essencial e tinha papel central na produção, que é uma das principais razões pelas quais o progresso levava a aumentos salariais. No entanto, se a IA e a robótica puderem substituir completamente qualquer tipo de profissão humana no futuro, ele deixará de ser essencial, e pode marcar o início do fim da era do trabalho. Por outro lado, se o fator escasso, que é o emprego, não for mais essencial para a produção, o crescimento econômico pode aumentar significativamente (KORINEK & JUELFIS, 2022).

Quando essa condição é satisfeita, o trabalho deixa de ser um insumo essencial para a produção, que pode ocorrer em qualquer escala desejada condicionada pelas máquinas substitutas. Assim, avanços em IA ocorrem a partir do desenvolvimento da computação e de seus algoritmos, e nos últimos anos, ambos exibem ritmos acelerados de crescimento. Mediante à potência de *hardware* atual, a capacidade de *software* das IA parecem ficar para trás até o momento. Entretanto, o mercado está atraindo um influxo significativo de talento e recursos nessa área, indicando que o progresso continuará sem interrupções podendo replicar melhor a função humana (KORINEK & JUELFIS, 2022).

A substituição perfeita não necessariamente implica redundância econômica do trabalho, uma vez que não considera o quão custoso é substituir o emprego de indivíduos por máquinas. Na hipótese de ser tecnicamente possível fazer essa substituição, mas isso custar muitas vezes mais do que os salários de mercado atuais dos empregados, então não seria economicamente eficiente fazê-lo. Assim, o custo pelo qual as máquinas podem desempenhar um determinado trabalho humano define o limite da substituição (KORINEK & JUELFIS, 2022).

À medida que a tecnologia avança e as máquinas se tornam cada vez mais eficientes, esse limite, e por implicação os salários de mercado, tendem a diminuir. Inicialmente, empregos que são facilmente automatizados são afetados, levando os humanos a mudarem para trabalhos mais difíceis de automatizar e, portanto, com salários mais altos, como têm feito há séculos. No entanto, no passado, sempre havia empregos que apenas os humanos poderiam realizar, que não será mais o caso se as máquinas se tornarem substitutos perfeitos para o trabalho. E à medida que as máquinas continuam a se tornar mais eficientes, os salários podem cair abaixo do custo necessário para empregar um humano, de modo que o mesmo não seja mais economicamente viável (KORINEK & JUELFIS, 2022).

A terceira possibilidade é que o trabalho se torne economicamente redundante, quando os preços das máquinas que o substituem são menores do que o custo de emprego humano. Em um mundo assim, os seres humanos não poderiam mais sobreviver com base na obtenção de salários competitivos de mercado. O trabalho se tornaria obsoleto e deixaria de desempenhar o papel central que tem atualmente na sociedade. No entanto, o capitalismo precisa de consumo para funcionar, e nesse sistema o emprego é o fator que impulsiona a renda disponível para consumo. Sem ele, a renda teria que ser proveniente de outra fonte (KORINEK & JUELFS, 2022; FINQUELIEVICH, 2019).

Então, caso o trabalho se torne economicamente redundante, outras fontes de arrecadação serão essenciais para evitar uma miséria em massa e a instabilidade política que pode resultar disso. Isso pode ser alcançado a partir de distribuição de capital ou de pagamentos de benefícios, uma vez que as máquinas que tornam o trabalho redundante seriam capazes de produzir em níveis sem precedentes e, a princípio, seria possível realizar uma distribuição mais equitativa. Como a grande maioria da população precisará de apoio financeiro, uma Renda Básica Universal para todos pode ser viável (KORINEK & JUELFS, 2022).

Uma economia mais rica e desenvolvida não significa que todos se beneficiarão igualmente, ou que sequer vão receber benefícios. No entanto, nessa situação os efeitos de produtividade da IA estarão intrinsecamente ligados a uma significativa perturbação no mercado de trabalho, com desemprego em massa e pressões salariais negativas, reforçando uma maior necessidade de auxílio (BAILY & KORINEK, 2023).

As últimas duas possibilidades de substituição de mão-de-obra, tanto com um custo mais alto quanto mais baixo que a de um humano, refletem a maneira como a potencial futura redundância do trabalho é frequentemente abordada pelas pessoas. Nesse contexto, a troca generalizada de indivíduos por máquinas a preços menores não aparenta ser razoável no cenário atual. No entanto, só de existir a possibilidade de substituição a custos maiores, junto a um nível de *software* e *hardware* que evoluem constantemente, isso implica na possibilidade dessas tecnologias se tornarem viáveis eventualmente, tornando esta questão uma contínua preocupação para a economia (KORINEK & JUELFS, 2022).

Nesse caso mais extremo, a humanidade não adicionaria mais valor econômico, o emprego humano custaria mais recursos do que seriam capazes de produzir, e se tornaria obsoleto, sendo dominado pela tecnologia. Incapaz de se sustentarem, a humanidade teria de escolher entre uma miséria em massa ou fornecer uma renda básica de subsistência, que podem

ser caras em um mundo dependente do trabalho humano, mas acessíveis em um mundo de prosperidade impulsionada pela IA (KORINEK & JUELFS, 2022; LITTMAN *et al.*, 2021).

Não existem leis físicas ou econômicas que sugiram que a inteligência e destreza das máquinas não possam, em princípio, ultrapassar as capacidades humanas. A inteligência humana está sujeita a limites naturais significativos, que por serem biológicos, são difíceis de superar. Com isso, os avanços dos algoritmos e máquinas têm excedido em muito os avanços na inteligência humana nas últimas décadas, que nesse ritmo eventualmente será ultrapassada. Em trabalhos mecânicos ou de menor habilidade cognitiva, a IA já consegue ser mais prática e barata do que os humanos (KORINEK & JUELFS, 2022; FINQUELIEVICH, 2019).

Além disso, para as possibilidades de substituição do trabalho, tanto com um custo mais alto quanto mais baixo que o de um humano, as máquinas não precisam necessariamente ter algo correspondente à consciência humana. De fato, elas precisam ser apenas capazes de realizar tarefas que têm valor econômico de maneira eficaz. Nesse caso, inclui tarefas que envolvem inteligência social e emocional, o que exigirá que teorias suficientemente avançadas da mente sejam desenvolvidas, ainda não reproduzidas pela IA (KORINEK & JUELFS, 2022).

Já em relação à dinâmica populacional, uma possibilidade é que a IA gere impactos diretos na relação entre a demografia e economia. As populações estão envelhecendo em todo o mundo, e em alguns países desenvolvidos, as forças de trabalho já estão diminuindo. Pode ser que a IA ajude a amenizar a escassez de mão-de-obra decorrente do aumento de aposentados, pois a produtividade gerada pelos algoritmos compensaria pela queda na oferta de profissionais, que por consequência implicariam déficit nas produções (LITTMAN *et al.*, 2021).

Assim, em países onde a população em idade ativa está diminuindo, o impacto da automação de empregos na força de trabalho seria amenizado, uma vez que os trabalhadores substituídos pela IA também estão mais perto de se aposentar. Além disso, as populações envelhecidas geram uma grande demanda por serviços, especialmente na área da saúde, muitos desses empregos no setor de serviços provavelmente não serão substituídos em um futuro próximo (KORINEK & STIGLITZ, 2021).

5.4 Potencial da Inteligência Artificial generativa

Até recentemente, trabalhos com interação, como o atendimento ao cliente, têm recebido intervenções tecnológicas menos maduras, mas a IA generativa está prestes a mudar isso. A IA, de uma maneira que se aproxima e, em alguns casos, é indistinguível do comportamento humano, pode assumir esse processo de interação. Isso não significa que essas

ferramentas são projetadas para funcionar sem a contribuição e intervenção humanas e, em muitos casos, são mais poderosas quando usadas por eles ao invés do seu próprio sistema automático, aumentando suas capacidades e permitindo que realizem o trabalho de forma mais rápida e eficiente (CHUI & ROBERTS & YEE, 2022).

Esses modelos estão nos estágios iniciais de ampliação, mas já é perceptível a possibilidade dos usos de vários modelos de aplicação em diversas funções e áreas do trabalho. Em sua maioria as empresas podem adotar essa tecnologia para ajudar os trabalhadores no processo criativo, visto que suas capacidades autônomas sem supervisão ou correções ainda são limitadas e muitas vezes insatisfatórias, por enquanto. Alguns exemplos do uso de IA generativa como ferramentas para aumentar a eficácia do trabalho são:

- Marketing e vendas – criação de conteúdo de marketing personalizado, mídias sociais e conteúdo técnico de vendas (incluindo texto, imagens e vídeo); e criação de assistentes específicos para empresas, como varejo;
- Operações – geração de listas de tarefas para a execução eficiente de uma determinada atividade;
- TI/engenharia – escrita, documentação e revisão de códigos;
- Jurídico – resposta a perguntas complexas, consulta a vasta quantidade de documentação jurídica, redação e revisão de relatórios (CHUI & ROBERTS & YEE, 2022).

No entanto, com o desenvolvimento das capacidades generativas dessa tecnologia, eventualmente a mesma pode deixar de auxiliar em tarefas e assumir processos inteiros, chegando até a substituição de empregos. A princípio o impacto deve ser maior em trabalhos repetitivos, rotineiros e previsíveis por natureza, como em áreas administrativas. Porém, a longo prazo, a IA pode atingir empregos que exigem inteligência humana e alto nível de educação, como escrita, jornalismo, tradução, direito e programação. Um estudo recente mostra que os LLMs podem afetar 80% da força de trabalho dos Estados Unidos, desde tarefas até profissões (BAILY & KORINEK, 2023; AHN & CHEN, 2023).

Com essa substituição, mesmo que parcial, menos pessoas serão necessárias para produzir a mesma quantidade de informação, desequilibrando a demanda de profissões no mercado de trabalho. É improvável que isso substitua completamente agentes humanos, mas considerando a sofisticação de sistemas como ChatGPT, é provável que essa redução seja significativa em um futuro próximo. Em vez de contratar dez programadores, podem ser necessários apenas dois programadores que possam revisar, corrigir e editar o código criado pela IA (AHN & CHEN, 2023).

Historicamente, mesmo quando novas tecnologias substituíam empregos, aqueles que escreviam as instruções para as máquinas se sentiam imunes. No entanto, recentemente, muitos trabalhadores dessa área já expressam crescente ansiedade em relação à súbita ascensão da IA generativa. Aqueles que estão automatizando processos temem que em breve possam ser substituídos pela própria ferramenta (TURNER, 2023).

Alguns trabalhadores já estão utilizando a ferramenta em suas profissões, para escrever descrições de imóveis para agentes imobiliários, descrições de vagas de emprego para gerentes de contratação e discursos de vendas para profissionais de marketing. Ele se destaca em tarefas de redação padronizadas, porque imita textos que já existem *online*, fazendo de forma fácil os primeiros rascunhos deste tipo de tarefa. O ChatGPT resolve o problema de ter que começar um documento do zero ao fornecer um texto base ao profissional. Dessa forma, ele promove uma ajuda inicial, mas por enquanto ainda é uma ferramenta limitada e com falhas, ou seja, não consegue substituir o trabalho, apenas uma tarefa presente nele (JACKSON, 2023).

A IA continuará a transformar indústrias à medida que as empresas a integram em uma variedade de tarefas, como as citadas acima. O ChatGPT provavelmente eliminará algumas funções, mas já cria novos empregos: algumas companhias estão contratando pessoas para testar e vender as consultas mais eficazes do *chatbot*. Como a IA é projetada para aprender, ela só vai melhorar, e em algum momento os *chatbots* poderão escrever produtos finais em vez de apenas rascunhos, e eliminar muito mais empregos do que conseguem gerar, mas até lá, a maioria deles está segura (JACKSON, 2023).

Diversas literaturas emergentes estimam ganhos de produtividade por IA em diferentes ocupações. Uma pesquisa mostra que engenheiros de *software* podem codificar duas vezes mais rápido, escritores podem concluir tarefas em menor tempo e economistas podem ser 10 – 20% mais produtivos usando LLM. Além disso, operadores de telemarketing se tornaram 14% mais produtivos com a ajuda da IA, que resultaram em maior satisfação do cliente e menor rotatividade de funcionários, e com isso um aumento de 30% na receita. Esses avanços indicam que a IA pode trazer ganhos significativos com o aumento de produtividade, capturando e transmitindo conhecimentos que só eram obtidos com a experiência diária no trabalho (BAILY & KORINEK, 2023).

5.5 Setores de tecnologia impulsionados pela Inteligência Artificial generativa

O progresso da IA generativa está fortemente ligado ao poder de processamento atual, como resultado, grande parte do dinheiro acaba fluindo para companhias de infraestrutura

computacional. Foi estimado que, em média, as empresas de IA gastam cerca de 20 a 40% da receita para o funcionamento do algoritmo e nos ajustes específicos para cada cliente. Isso é geralmente pago diretamente a provedores de nuvem ou a fornecedores terceirizados, que por sua vez, gastam cerca de metade de sua receita em infraestrutura de nuvem. Portanto, é previsto que de 10 a 20% da receita total em IA generativa hoje seja destinada aos provedores de nuvem de firmas líderes no setor de TI (BORNSTEIN & APPENZELLER & CASADO, 2023).

Nesse sentido, muitas empresas de tecnologia, incluindo muitas *startups*, gastam milhões de dólares por ano no treinamento de IA, seja com provedores externos de nuvem ou diretamente com fabricantes de *hardware*. Assim, devido ao alto custo de inovar, esse mercado de IA movimentou quantidades enormes de capital já em sua fase inicial. A maior parte é gasta nos três principais provedores de nuvem: Amazon Web Services (AWS), Google Cloud Platform (GCP) e Microsoft Azure. Esses provedores de nuvem despendem coletivamente mais de 100 bilhões de dólares por ano em despesas para garantir que tenham as plataformas mais abrangentes, confiáveis e competitivas em termos de custo (BORNSTEIN & APPENZELLER & CASADO, 2023).

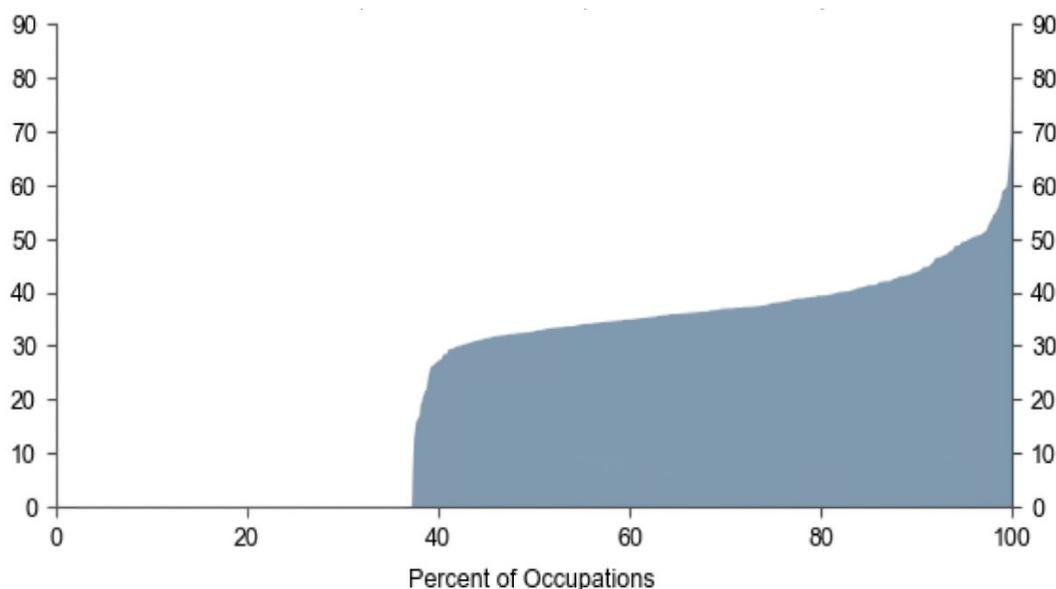
Algumas corporações também se beneficiam da oferta limitada de *hardware*, já que para ter o mínimo de competitividade nesse meio, o básico é se manter atualizado com as peças mais novas disponíveis no mercado, no entanto, elas são escassas, como as GPUs Nvidia A100 e H100 já mencionadas. Como resultado, fornecendo a grande maioria dessa tecnologia, está talvez a maior vencedora até agora nesse mercado, a Nvidia, que só no primeiro trimestre de 2023 teve uma receita de US\$ 3,8 bilhões com GPUs. Por dominar esse mercado, por meio de décadas de investimento na arquitetura de GPUs, formaram-se fortes barreiras de entrada para esse segmento altamente importante no contexto atual de inovações por IA. Uma análise recente mostra que as GPUs Nvidia são citadas em artigos de pesquisa 90 vezes mais do que as principais *startups* de *chips* de IA combinadas (BORNSTEIN & APPENZELLER & CASADO, 2023).

5.6 Impacto do nível de trabalho pela Inteligência Artificial generativa

Com as possibilidades de uma aceleração rápida na automação pela nova capacidade generativa da IA, se ela atender às expectativas previstas, o mercado de trabalho poderá enfrentar uma significativa perturbação. Utilizando dados sobre profissões nos EUA e UE, foi estimado que cerca de dois terços dos empregos atuais estão expostos a algum grau de automação, e que a IA generativa poderia substituir até um quarto da mão-de-obra vigente.

Projetando essas estimativas para o âmbito global, é possível inferir que essa tecnologia poderia sujeitar cerca de 300 milhões de empregos à automação (HATZIUS *et al.*, 2023).

Gráfico 14 – Nível percentual da distribuição da carga de trabalho ocupacional exposta à automação por IA, 2023



Fonte: HATZIUS *et al.*, 2023;

Nesse estudo, representado pelo gráfico 14, foram usadas estimativas da participação do emprego de cada ocupação na Pesquisa de Emprego e Salário Ocupacional dos Estados Unidos e depois agregadas ao nível da indústria. Desses dois terços de ocupações expostas em algum grau de automação, a maioria tem uma parcela significativa de sua carga de trabalho, de 25% a 50%, que pode ser substituída. Já o um quarto possível de ser substituído tem exposições particularmente altas nas profissões administrativas (46%) e jurídicas (44%) e baixas em profissões fisicamente intensivas, como construção (6%) e manutenção (4%). Ao combinar essas estimativas com o sistema de classificação de ocupações (ISCO) da Europa e realizar um exercício semelhante para essa área usando o banco de dados da Pesquisa da Força de Trabalho do Eurostat (LFS), chegou-se a estimativas semelhante, tanto em termos agregados quanto por setor (HATZIUS *et al.*, 2023).

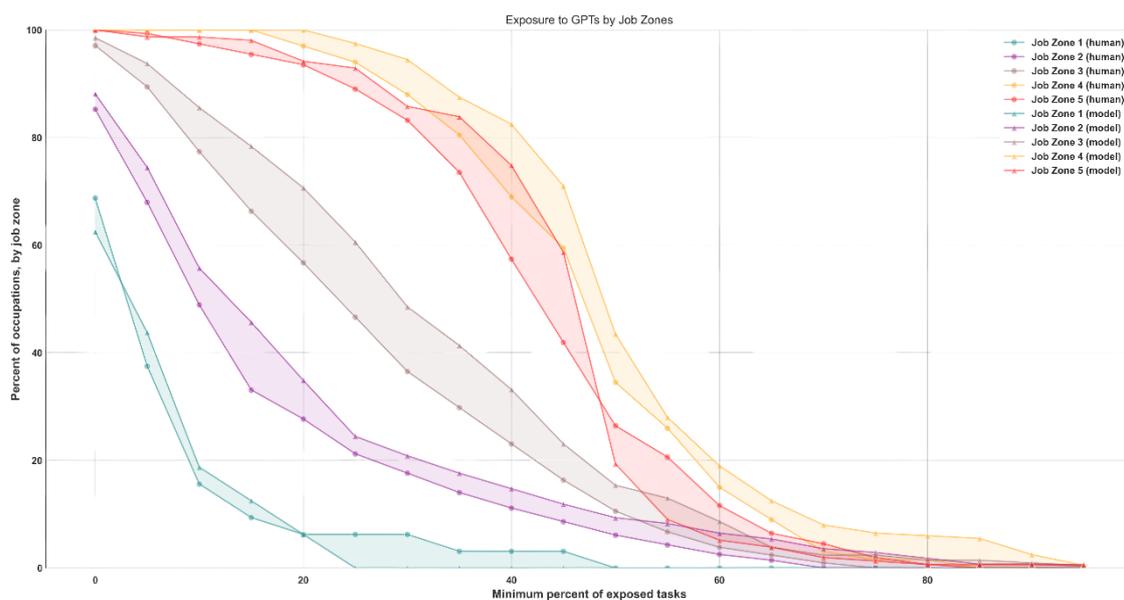
Visto que as IA generativas afetarão os trabalhadores que utilizam processos criativos, diferente de outros pontos da história em que a automação atingia somente os trabalhos menos qualificados, os empregos com salários altos também serão impactados. Um estudo realizado pela OpenAI, OpenResearch e UPenn mostra o potencial dos LLMs de afetar várias ocupações

e setores da economia dos Estados Unidos (FELTEN & RAJ & SEAMANS, 2023; ELOUNDOU *et al.*, 2023).

Este estudo mostra as classificações de exposição do modelo de IA em cinco níveis, representado no gráfico 15, formados por grupos de ocupações semelhantes classificadas de acordo com o nível de educação, experiência e habilidade necessária para realizá-la. O resultado aponta que competências que envolvem ciência e pensamento crítico têm menos probabilidade de serem impactadas pelos modelos atuais. Por outro lado, as tarefas de escrita e programação são mais suscetíveis a serem influenciadas. Além disso, indivíduos com diploma de Bacharelado, Mestrado e cursos profissionalizantes estão mais expostos a IA generativa e *softwares* com LLM do que aqueles sem certificações educacionais formais, mas que adquiriram habilidades por experiência no trabalho (ELOUNDOU *et al.*, 2023).

Considerando as capacidades atuais do modelo, os resultados apresentados no gráfico 15 mostram que cerca de 80% da força de trabalho dos EUA poderia ter pelo menos 10% de suas tarefas afetadas, enquanto aproximadamente 19% dos empregados chegam a 50% de suas tarefas influenciadas. Embora a capacidade da IA de tornar o trabalho humano mais eficiente seja evidente, é importante reconhecer que fatores sociais, econômicos, regulatórios e outros influenciarão nos resultados reais do ganho de produtividade. À medida que as capacidades continuarem a evoluir, o impacto dos LLMs na economia aumentará, e com isso os desafios gerados por eles (ELOUNDOU *et al.*, 2023).

Gráfico 15 – Nível percentual de exposição a tecnologia GPT por grupo de emprego



Fonte: Eloundou *et al* (2023)

O grupo de emprego 1, em azul, requer a menor quantidade de preparação, 3 meses, e o grupo 5, em vermelho, demanda uma maior quantidade, de 4 anos ou mais. A renda mediana aumenta de forma constante entre os grupos de emprego, à medida que o nível de preparação necessário também aumenta, sendo que o trabalhador mediano do primeiro nível ganha US\$ 30.230 e o do quinto nível ganha US\$ 80.980, por ano nos EUA. Existe um certo padrão entre os grupos, sendo que a exposição à IA aumenta da zona ocupacional 1 até a 4 e 5, possivelmente afetando os trabalhadores mais qualificados (ELOUNDOU *et al.*, 2023).

5.7 Variação do nível de emprego atual

Analisando dados históricos, a inovação que inicialmente substituiu trabalhadores impulsionou o crescimento da mão-de-obra ao longo prazo. Um estudo relata que 60% dos trabalhadores de hoje estão empregados em ocupações que não existiam em 1940. Isso implicou que mais de 85% do crescimento do emprego nos últimos 80 anos ocorreu no âmbito de novas posições impulsionada pelo avanço tecnológico. É válido considerar que, por mais que a natureza da IA seja diferente das inovações passadas pelo seu poder generativo, o cenário futuro é incerto, logo o mais prudente a se fazer é avaliar as possibilidades condizentes com esse progresso (HATZIUS *et al.*, 2023).

Mesmo com os diferentes cenários, o desemprego inclusive de profissionais mais qualificados não é hipotético, pois ele já é uma realidade em algumas corporações. Por exemplo, na instituição financeira ING, a implementação de aplicativos de IA na área de regulatória de *compliance* resultou na substituição de 400 funcionários em tempo integral. Foi relatado que a empresa ensinou o *software* a lidar com tarefas básicas de escritórios que os trabalhadores estavam fazendo, com isso a produtividade da companhia aumentou e os custos diminuíram com o corte de despesas (RANSBOTHAM *et al.*, 2018).

O presidente executivo da IBM, Arvind Krishna, afirmou que considera fazer uma redução drástica de seu quadro administrativo, dado o potencial da IA para executar tarefas nesta área. Assim a companhia está paralisando contratações como parte de um plano para implementar essa substituição. Cerca de 26 mil pessoas trabalham em postos que não lidam com clientes e segundo o CEO, uma parcela desse número (7.800) pode ser trocada por IA em cinco anos (PRESSE, 2023; TURNER, 2023).

Corporações como Meta, Microsoft e Salesforce têm despedido funcionários nos últimos meses, a justificativa é a busca por maior eficiência, redução de custos e aumento das margens de lucro. De acordo com estimativas, mais de 250.000 trabalhadores de tecnologia

foram demitidos desde o início de 2022. Pelo fato do futuro ser imprevisível, além de despedir pessoas, muitas empresas estão suspendendo as contratações e a procura de algumas áreas profissionais dentro da TI – profissionais da área de IA são exceção a isso. No entanto, afirmar que o motivo atrelado a esse processo é a difusão da IA é incoerente, visto que existem fatores mais consolidados que a IA a serem levados em conta (TURNER, 2023).

Gráfico 16 – Número de demissões na área de TI por trimestre, 2020 – 2023



Fonte: Layoffs.fyi 2023

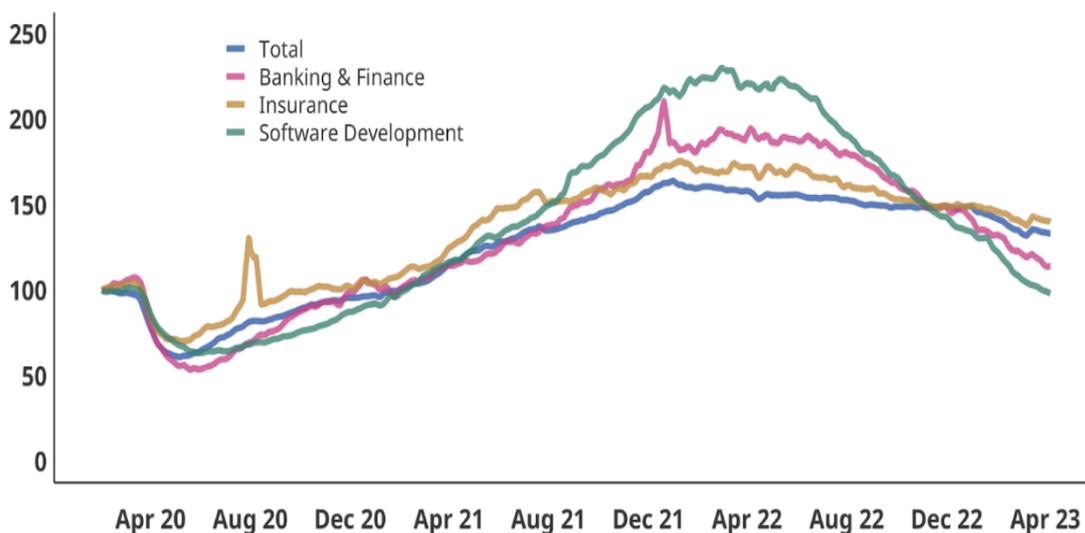
Como pode ser visto no gráfico 16 do *site* Layoffs.fyi, empresas de tecnologia demitiram um total de 164.709 funcionários ao longo de 2022 de acordo, e no ano de 2023, até maio, já foram noticiados o corte de 201.566 empregos no mundo. Os picos de despedidas ocorreram no começo da pandemia e na transição de 2022 para 2023. O *site* registra e contabiliza esse processo de demissões feitas e anunciadas para o período analisado, vindo de declarações, artigos de notícias e anúncios de companhias de todo o mundo, em tempo real e organizados por data. A lista é extensa e provavelmente não chega perto de capturar o total de perda de empregos que realmente acontece, alguns exemplos são:

- A Alphabet tem planos de despedir 12 mil empregados, que representa quase 6% do seu total de trabalhadores em 2023;
- A Meta despediu mais de 11 mil empregados, 13% do seu total de trabalhadores, e segundo a companhia, eles entraram em um congelamento de contratações pelo período inicial de 2023;
- A Amazon declarou que começando em novembro de 2022, irá cortar em torno de 18 mil funcionários de sua empresa;

- A Microsoft anunciou que demitirá em torno de 10 mil funcionários durante 2023, afetando em torno de 5% da força de trabalho da companhia;
- A Dell relatou um corte de 6.650 trabalhadores, que representa em torno de 5% da força de trabalho da companhia, chegando a o número mais baixo desde 2017;
- A Yahoo dispensou mais de 20% de sua equipe ao reduzir seu negócio de publicidade, os cortes de empregos afetarão mais de 50% dos funcionários de tecnologia de anúncios da empresa, estimados em mais de 1.600 pessoas;
- A Twilio anunciou a redução de 17% de sua força de trabalho, equivalente a cerca de 1.500 empregados. A companhia de *software* de comunicações em nuvem já havia demitido 11% de seus funcionários antes deste anúncio;
- A rede social LinkedIn, de propriedade da Microsoft, demitiu 716 funcionários pela descontinuidade de um de seus projetos relacionado a China;
- A Ford, que está passando por uma transição para veículos elétricos, está demitindo 3.800 funcionários nas áreas de engenharia e administração, o que indica possíveis demissões adicionais nos setores automotivo e de veículos elétricos (LAYOFFS.FYI, 2023; SEVILLA, 2023; CHOW, 2023).

Em relação à queda na oferta de postos de trabalhos, embora o *site* Indeed reportar uma redução nas mesmas, elas estão bem acima dos níveis pré-pandemia. Além disso, as demissões no mercado como um todo permanecem historicamente baixas, e a rotatividade de funcionários diminuiu na indústria de serviços. Como apresentado no gráfico 17, no primeiro trimestre de 2023, a baixa na oferta de empregos continuou refletindo a recessão do mercado de trabalho, que foi 15% abaixo do mesmo período do ano anterior. No entanto, a demanda por trabalhadores ainda está bastante forte, nesse mesmo período, pois estão 33% acima da fase pré-pandemia (CULBERTSON, 2023).

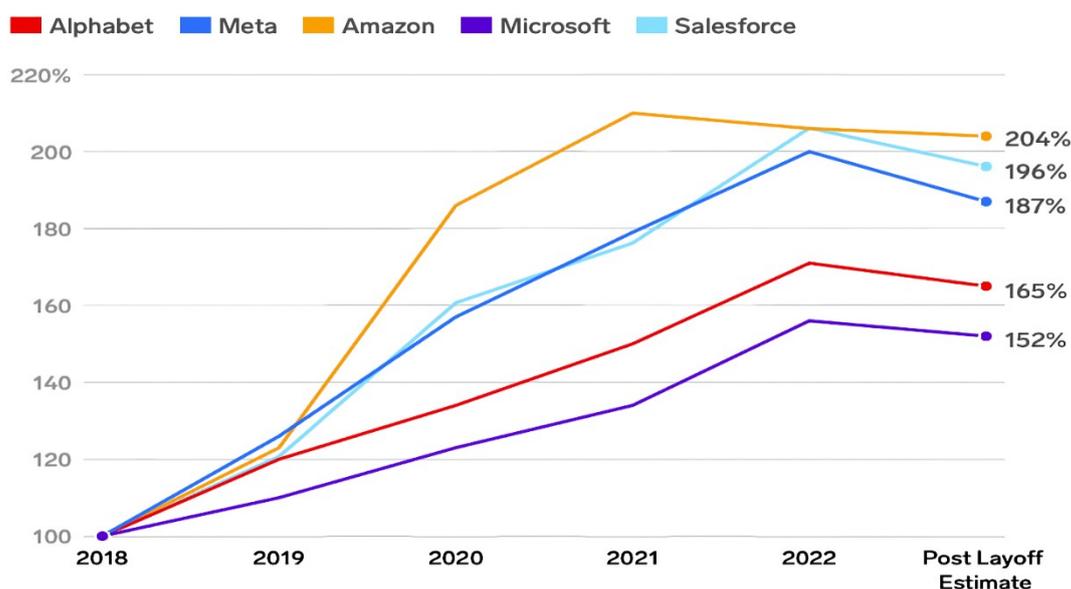
Gráfico 17 – Índice de postagens de emprego no Indeed de fev. 2020 – abr. 2023



Fonte: Indeed; CULBERTSON, 2023.

As postagens de empregos para segmentos específicos continuaram a desacelerar juntamente com o mercado de trabalho em geral, apresentado no gráfico 17. A queda mais significativa reflete o recuo no setor de tecnologia, com as postagens de cargos de desenvolvimento de *software* reduzidas em 55,4% em comparação com o mesmo período de 2022, enquanto isso, no setor de Bancos e Finanças diminuíram 41% e 18,5% na área de seguros (CULBERTSON, 2023).

Gráfico 18 – Empregos nas grandes empresas de tecnologia antes e depois das ondas de cortes de gastos, 2018 – 2023



Fonte: Chay; Insider.

O gráfico 19 revela que, apesar do aumento de demissões entre as principais empresas de tecnologia nos últimos meses, mesmo com os cortes, elas ainda mantêm um contingente maior de funcionários em comparação ao período pré-pandemia. Em outras palavras, apesar das tendências de redução de custos e demissões em massa, as grandes firmas de tecnologia, de maneira geral, registraram um aumento significativo no número de funcionários ao longo dos últimos anos. Esse padrão reflete uma tendência de contratação que teve início muito antes da pandemia impulsionar esse processo (ZAVERI, 2023).

Nesse sentido, as análises indicam que essas demissões têm muito mais a ver com as condições econômicas atuais do que com avanços recentes em IA. O surgimento da IA generativa ainda pode ter um impacto enorme na força de trabalho, mas, por mais que a tecnologia mostrou seu altíssimo potencial, esse processo ainda leva tempo para ser perceptível. Esses cortes podem ser inicialmente associados à expansão de TI durante a pandemia, à medida que as pessoas passaram muito mais tempo em seus computadores domésticos, a demanda e oferta por serviços *online* disparou. Como resultado, as empresas de tecnologia cresceram rapidamente e obtiveram lucros recordes. A Microsoft foi uma das mais beneficiadas, com lucros crescendo exponencialmente em 2021, assim, entre junho de 2021 e 2022, eles contrataram 40.000 funcionários, número recorde da companhia (CHOW, 2023).

A questão é que durante a pandemia, quando as taxas de juros estavam muito baixas, as empresas tinham muito dinheiro à disposição, sendo que os investidores priorizavam no crescimento, e não na lucratividade. No entanto, à medida que as taxas de juros estão subindo – que foi a realidade após a pandemia, e como foi anunciado para o ano de 2023 segundo o Sistema de Reserva Federal dos EUA – houve uma mudança por parte dos grandes investidores para inverter esse processo. Logo, o foco mudou para estratégias que potencializem a lucratividade, e uma das primeiras maneiras de fazer isso é por meio de cortes de gastos, que resultam em demissões generalizadas (CHOW, 2023; BENANAV, 2023).

Em relação à busca por produtividade, e garantir ou diminuir, qualquer vantagem de mercado em relação à concorrência, as companhias tomam suas decisões individuais buscando o resultado mais benéficos para elas. Logo, contando que a IA atenda seus requisitos de qualidade, caso possam aumentar a produtividade substituindo pessoas por máquinas, as empresas o farão. Assim como as corporações têm cortado empregos nos últimos meses, muitas vezes em busca de eficiência e aumento das margens de lucro, elas farão o mesmo quando a IA for mais viável, que ainda não é o caso (TURNER, 2023).

Em relação a perspectivas sobre a substituição, Satya Nadella, CEO da Microsoft, declarou que, a empresa planejava incorporar a IA em todos os seus produtos, e que os trabalhadores se tornariam mais eficientes e produtivos em conjunto com a IA ajudando a fazerem mais com menos, implicando que essas tecnologias atuariam como um copiloto, e não substituto. Já na parte da criação de empregos, o Fórum Econômico Mundial fez uma previsão que a IA substituiria 85 milhões de trabalhadores até 2025, mas criaria novos 97 milhões (CHOW, 2023).

Mediante a análise do cenário atual em relação ao nível de emprego, e levando em conta as teorias sobre a possibilidade da substituição de trabalho por IA, existem poucos sinais de que esse processo tenha se desenvolvido. De fato, há casos onde ela já desempenha tarefas antes feitas por humanos, mas isso não difere o processo da IA de outras inovações menos disruptivas. Diante de posições e previsões, tanto otimistas quanto pessimistas, o fato é que o cenário da IA ainda é incerto e prematuro. A adesão generalizada dela ainda está longe de ser uma realidade, é preciso mais tempo para entender como as previsões vão se manter perante ao evidente potencial da mesma.

CONSIDERAÇÕES FINAIS

O presente trabalho investigou os possíveis impactos da IA na economia e sociedade de forma ampla, avaliando a forma que ela modifica as técnicas e percepções dos indivíduos nesse meio. Com isso, foram decorridos o processo de difusão da tecnologia, seus riscos e desafios éticos, a sua regulação, a IA generativa e, por fim, seu potencial disruptivo no mercado de trabalho.

No que se refere à sua difusão, foi constatado que seus impactos variam de acordo com inúmeros fatores, como a área de uso, aplicabilidade, investimento feito, nível de intensidade, entre outros. Nesse sentido, constatou-se que os setores com maior destaque em todos esses quesitos foram os setores de maior intensidade tecnológica, especialmente a própria computação. Pode-se concluir que o motivo disso é por serem os criadores dessas tecnologias e estarem à frente do processo de impulsão das mesmas. Além disso, até o momento, por obterem os melhores resultados financeiros com a IA, e pela compreensão que têm sobre a mesma, elas são mais dispostas a arriscar com essa mudança tecnológica.

No que tange à utilização da tecnologia, é perceptível que a implementação de IA em diversos setores de economia foi mais constante e intensiva nos últimos anos devido ao seu desenvolvimento. Não exclusivo da IA, mas fica evidente que existe uma defasagem entre o surgimento de uma nova tecnologia e seu uso em escala. Nesse sentido, foi possível perceber que sua complexidade prejudica o processo de compreensão das pessoas em relação a seu uso e implicações, e por consequência a sua difusão.

Em relação à localidade, verificou-se que a intensidade da exposição de IA depende do nível de desenvolvimento do país. Logo, seu uso em nações emergentes é menor. Nesse sentido, a América Latina possui uma baixa difusão da tecnologia, mas isso possibilita muitas oportunidades para o uso da IA. Já na perspectiva dos líderes, a China, EUA e UE se colocam muito distante de quaisquer outros países em todos os quesitos.

Em relação à ética na IA, por se tratar de uma ferramenta, foram abordados vários casos de usos benéficos e prejudiciais da tecnologia, e cabe ao usuário definir qual será a utilidade da mesma. Dessa forma, se mostra uma tarefa árdua garantir que ela exerça apenas os fins realmente planejados e não gere consequências indesejadas, transgredindo barreiras éticas e legais. Assim, mesmo quando os criadores do algoritmo não têm a intenção de gerar malefícios, eles devem ser responsáveis por prevenir e lidar com tais problemas.

Políticas e regulações dos governos são claramente uma parte essencial do processo de inovação. Entretanto, o ritmo acelerado do avanço tecnológico coloca os governos em uma desvantagem considerável. Assim, constatou-se que seria muito benéfico uma cooperação internacional, para garantir que a IA seja usada em prol da sociedade, sem comprometer a jurisdição e qualidade de vida de outros países. No entanto, verificou-se que existe uma disputa entre os líderes em IA, em que os EUA e UE declararam sua oposição às medidas autoritárias tomadas pela China.

A UE foi pioneira na regulamentação e governança de IA, inspirando outras leis tanto nos EUA como na China. Porém, conclui-se que seu objetivo implícito com o AI Act é estimular a competitividade europeia, estabelecendo um padrão global de regulação para gerar uma conformidade internacional. Assim, fabricantes estrangeiros que se ajustarem às regras europeias, uma vez feito gasto com adaptação, terão um forte incentivo para continuar a cumprir essas regras.

A abordagem dos EUA em relação à IA é movida pelo ideal de que as empresas, em geral, devem permanecer no controle do desenvolvimento e critérios relacionados à governança. Constatou-se que na visão deles, intervenções e regras excessivas representam barreiras à inovação. Devido à capacitação do país em IA, os princípios estabelecidos no AI Bill of Rights poderiam ser eficientes para mediar o uso da IA. No entanto, como não são vinculados às leis, o grau de mudanças depende dos governos de cada estado, que por consequência atrasam o processo de governança.

Por outro lado, na China, as leis permitem que o Estado exerça controle sobre as empresas privadas, e as guiem de acordo com seus interesses, exigindo que sigam em comum com as ideias do Partido Comunista. Com isso, pode-se concluir que há uma ampliação na divergência das suas práticas com a dos mercados internacionais.

Em relação à América Latina, poucos países lançaram estratégias autônomas relacionadas à IA. Devido à baixa capacitação das nações dessa região no setor tecnológico, constatou-se que os planos lançados têm como intuito muito mais se desenvolver uma maior qualificação nesta área do que regular algo que mal se faz presente na região. Verificou-se a falta de conhecimento, experiência e expertise, que faz com que a adoção de regulações seja condizente com sua realidade, exigindo mais prudência no processo de adesão à tecnologia.

Já no contexto da IA generativa, assumindo um papel diferente na IA, a tecnologia possibilita a criação de conteúdos próprios. Identificou-se que ainda existem limitações, mas a facilidade dos usuários de interagirem com os sistemas por meio da linguagem natural humana,

permite maior acessibilidade à IA do que modelos anteriores. Porém, foi possível estabelecer que a novidade também tem problemas, como a geração de conteúdos errados, parciais, perigosos ou *Fake News*. Além disso, constatou-se que a construção desses modelos possui um custo elevado devido a sua complexidade, ao ponto que apenas algumas grandes companhias da tecnologia chegaram a resultados mais notórios até o presente.

As empresas que mais se beneficiam nesse meio, até o momento, são a OpenAI, desenvolvedora do ChatGPT e do DALL-E, a DeepMind, subsidiária da Alphabet e a Meta com o Make-A-Video. Além disso, a maior parte dos gastos com essas IAs favorecem três principais provedores de nuvem: Amazon Web Services, Google Cloud Platform e Microsoft Azure. Já no quesito *hardware*, por dominar o mercado de GPUs, os lucros são apropriados a Nvidia. No entanto, como os custos da implementação dessas tecnologias em massa ainda são caros, concluiu-se que a demanda computacional em escala não seria sustentável.

Em relação aos possíveis cenários disruptivos, percebeu-se que as inovações anteriores eram mais restritas às atividades mecânicas, mas a IA já consegue até realizar o processo criativo. Assim, o progresso tecnológico pode substituir mão-de-obra e reduzir os salários. E caso a tecnologia se torne mais eficiente, o custo pelo qual as máquinas podem desempenhar o trabalho humano definirão o limite dessa substituição.

No que tange à diferença de mão-de-obra entre os países, os efeitos de substituição ocorrem de forma mais intensa para os trabalhadores com menores níveis de educação, mais presentes em economias emergentes. O progresso tecnológico tende a perpetuar a dominação das empresas mais capacitadas a liderar esse processo de inovação, assim como deteriorar os termos de comércio internacionais para nações em desenvolvimento. Isso acontece por consequência das inovações ocorrerem nas economias desenvolvidas, fazendo com que os países em desenvolvimento precisem importá-las para manter o mínimo de competitividade.

No tocante à área de tecnologia, verificou-se que uma baixa na contratação do mercado e alta nas demissões é reflexo da recessão do mercado de trabalho. Dessa forma, as análises indicam que essas variações no mercado são uma resposta das condições econômicas atuais, e não dos avanços recentes em IA. Além disso, constatou-se que ainda assim, o número total e demanda por trabalhadores ainda está acima do que era antes da pandemia.

Com isso, ao analisar o potencial vigente de substituição de trabalho pela IA, por mais que seu impacto seja inevitável e que seja perceptível seu uso inicial em várias áreas, seu fator disruptivo ainda é muito reduzido. Pode-se alegar que existe a tendência de em um futuro próximo o impacto da IA crescer consideravelmente, mas por enquanto ele não é tão evidente.

Assim, a automação por IA acontece e continua em andamento, mas não aparenta estar desrespeitando o ritmo que outras inovações apresentaram em outros fluxos tecnológicos. Tal fato pode muito bem mudar, visto que seu potencial de automação é o mais abrangente exibido até hoje.

Por fim, as evidências expostas sobre esse processo disruptivo são amplas, várias possibilidades podem acontecer. Na visão otimista, a substituição é seguida por uma realocação para novos empregos e os ganhos de produtividade geram prosperidade econômica. Na visão pessimista, a demissão em massa gera um caos social e político, sendo necessário distribuir uma renda básica como remediação. Porém, mediante ao processo histórico já visto, é mais provável ocorrer o intermediário entre esses extremos, com a disrupção gerando danos à economia, mas com possibilidades de adaptação, fomentados pelas necessidades de regulação e suporte governamental.

REFERÊNCIAS BIBLIOGRÁFICAS

- AFFONSO, C. **Google enfrenta ação dos EUA e ChatGPT: entenda tudo o que está em jogo.** Tilt UOL. 2023. Disponível em: https://www.uol.com.br/tilt/colunas/carlos-affonso-de-souza/2023/02/07/chatgpt-e-acao-antitruste-colocam-google-em-alerta-maximo.htm?utm_campaign=tilt&utm_content=hyperlink-texto&utm_medium=email&utm_source=newsletter. Acesso em: 18 mai. 2023.
- AHN, M; CHEN, Y. **Building guardrails for ChatGPT.** BROOKINGS. 2023. Disponível em: https://www.brookings.edu/blog/techtank/2023/02/07/building-guardrails-for-chatgpt/?utm_campaign=Center%20for%20Technology%20Innovation&utm_medium=email&utm_content=247081757&utm_source=hs_email. Acesso em: 18 mai. 2023.
- ANYOHA, R. **The History of Artificial Intelligence.** SITN HARVARD. 2017. Disponível em: <https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/>. Acesso em: 17 mai. 2023.
- ARBEX, P. **Os planos do Google para o Brasil: cabos submarinos, IA e...** Gil. BRAZIL JOURNAL. 2022. Disponível em: https://braziljournal.com/os-planos-do-google-para-o-brasil-cabos-submarinos-ia-e-gil/?utm_source=BJ+subscribers&utm_campaign=9da78bb106-EMAIL_CAMPAIGN_5_5_2019_20_17_COPY_01&utm_medium=email&utm_term=0_850f0f7afd-9da78bb106-427748885. Acesso em: 17 mai. 2023.
- AUTOR, D. *et al.* **AN INCLUSIVE FUTURE? TECHNOLOGY, NEW DYNAMICS, AND POLICY CHALLENGES.** BROOKINGS. 2022
- AWS AMAZON. **What Is GPT?.** AWS AMAZON. 2023. Disponível em: <https://aws.amazon.com/what-is/gpt/#:~:text=Generative%20Pre%2Dtrained%20Transformers%2C%20commonly,AI%20applications%20such%20as%20ChatGPT>. Acesso em: 18 mai. 2023.
- BAILY, M; KORINEK, A. **Machines of mind: The case for an AI-powered productivity boom.** BROOKINGS. 2023. Disponível em: https://www.brookings.edu/research/machines-of-mind-the-case-for-an-ai-powered-productivity-boom/?utm_campaign=Economic%20Studies&utm_medium=email&utm_content=258560189&utm_source=hs_email. Acesso em: 20 mai. 2023.
- BALAKRISHNAN, T. *et al.* **Global survey: The state of AI in 2020.** McKinsey & Company. 2020.
- BASS, D. **A aposta de 10 bilhões de dólares da Microsoft com o ChatGPT.** EXAME 25 fev. 23. 2023.

BCG. **CO2 AI by BCG and CDP to Launch the Tech Platform of Reference to Enable the Exchange of Scope 3 Data and Accelerate Decarbonization Within Ecosystems**. BCG. 2022. Disponível em: <https://www.bcg.com/press/10february2022-bcg-cdp-build-tech-platform-scope-3-data-accelerate-decarbonization>. Acesso em: 17 mai. 2023.

BEHNCKEN, K. **The world needs cybersecurity experts – Microsoft expands skilling effort with a focus on women**. Microsoft. 2023. Disponível em: https://blogs.microsoft.com/on-the-issues/2023/04/19/cybersecurity-skills-initiative-expansion-nonprofits/?utm_source=GZERO%20Daily&utm_campaign=c1f3017ded-EMAIL_CAMPAIGN_2023_04_26_11_03&utm_medium=email&utm_term=0_e605619869-c1f3017ded-%5BLIST_EMAIL_ID%5D. Acesso em: 17 mai. 2023.

BENANAV, A. **AUTOMATION AND THE FUTURE OF WORK – 1**. Part 1: New Left Review 119, Sept Oct. 2019.

BERGLIND, N; FADIA, A; ISHERWOOD, T. **The potential value of AI—and how governments could look to capture it**. McKinsey & Company. 2022. Disponível em: [https://www.mckinsey.com/industries/public-and-social-sector/our-insights/the-potential-value-of-ai-and-how-governments-could-look-to-capture-it?hlkid=ae4c207b8a074b8386164ff217dd94d6&hctky=2092446&hdpid=22f3b7a5-6ea1-4561-8fba-a52ef9055949#/.](https://www.mckinsey.com/industries/public-and-social-sector/our-insights/the-potential-value-of-ai-and-how-governments-could-look-to-capture-it?hlkid=ae4c207b8a074b8386164ff217dd94d6&hctky=2092446&hdpid=22f3b7a5-6ea1-4561-8fba-a52ef9055949#/) Acesso em: 17 mai. 2023.

BORDEN, K. *et al.* **The future is now: Unlocking the promise of AI in industrials**. Five proven use cases show how artificial intelligence is creating value for advanced industrial companies. MCKINSEY & COMPANY. 2022. Disponível em: <https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/the-future-is-now-unlocking-the-promise-of-ai-in-industrials?stcr=3E4B15D08A72437E87FEF5CA6DEC93E7&cid=other-eml-alt-mip-mck&hlkid=551e19badfd447d78d78156ca9013a93&hctky=2092446&hdpid=bc49112a-d89a-4f82-bd80-d354443b0f07>. Acesso em: 17 mai. 2023.

BORNSTEIN, M; APPENZELLER, G; CASADO, M. **Who Owns the Generative AI Platform?.** a16z. 2023. Disponível em: <https://a16z.com/2023/01/19/who-owns-the-generative-ai-platform/>. Acesso em: 18 mai. 2023.

BOSWELL, B. *et al.* **An AI power play: Fueling the next wave of innovation in the energy sector**. How Vistra Corp. is partnering with McKinsey to improve efficiency and reduce emissions by using AI. McKinsey & Company. 2022.

BRANDIT, J. *et al.* **SUCCEEDING IN THE AI COMPETITION WITH CHINA A STRATEGY FOR ACTION**. Brookings Global China Project. 2022.

BULLOCK, J; KORINEK, A. **@elonmusk and @twitter: The problem with social media is misaligned recommendation systems, not free speech**. BROOKINGS. 2022. Disponível em: https://www.brookings.edu/research/elonmusk-and-twitter-the-problem-with-social-media-is-misaligned-recommendation-systems-not-free-speech/?utm_campaign=Economic%20Studies&utm_medium=email&utm_content=214113277&utm_source=hs_email. Acesso em: 17 mai. 2023.

CAFFERATA, J. *et al.* **O cenário da Inteligência Artificial e Advanced Analytics na América Latina**. McKinsey & Company. Brazil journal. 2020.

CALISKAN, A; STEED, R. **Managing the risks of inevitably biased visual artificial intelligence systems**. BROOKINGS. 2022. Disponível em: https://www.brookings.edu/blog/techtank/2022/09/26/managing-the-risks-of-inevitably-biased-visual-artificial-intelligence-systems/?utm_campaign=Center%20for%20Technology%20Innovation&utm_medium=email&utm_content=228370696&utm_source=hs_email. Acesso em: 17 mai. 2023.

CANDELON, F. A. *et al.* **The CEO's Guide to the Generative AI Revolution**. BCG. 2023. Disponível em: https://www.bcg.com/publications/2023/ceo-guide-to-ai-revolution?utm_medium=Email&utm_source=esp&utm_campaign=bcgx&utm_description=featured_insights&utm_topic=genai&utm_geo=global&utm_content=202303_2&utm_usertoken=03e92cab95c2405fa1a25306c4afe84632925c1d. Acesso em: 18 mai. 2023.

CHEATHAM, B; JAVANMARDIAN, K; SAMANDARI, H. **Confronting the risks of artificial intelligence**. McKinsey & Company. 2019.

CHOW, A. **AI Isn't to Blame for Layoffs at Microsoft and Other Tech Companies**. TIME. 2023. Disponível em: <https://time.com/6248290/microsoft-layoffs-ai/>. Acesso em: 29 mai. 2023.

CHUI, M. *et al.* **The state of AI in 2022—and a half decade in review**. McKinsey & Company. 2022.

CHUI, M; ROBERTS, R; YEE, L. **Generative AI is here: How tools like ChatGPT could change your business**. McKinsey & Company. 2022. Disponível em: <https://www.mckinsey.com/capabilities/quantumblack/our-insights/generative-ai-is-here-how-tools-like-chatgpt-could-change-your-business?cid=other-eml-ofl-mip-mck&hlkid=c40f7f40912e4dcfa9a7916cbc38e1b4&hctky=2092446&hdpid=42989045-434a-40cd-ab7e-3d75ebf84ed8>. Acesso em: 18 mai. 2023.

COLUMBUS, L. **10 Ways AI Has The Potential To Improve Agriculture In 2021**. Forbes. 2021. Disponível em: <https://www.forbes.com/sites/louiscolumbus/2021/02/17/10-ways-ai-has-the-potential-to-improve-agriculture-in-2021/?sh=43748f347f3b>. Acesso em: 17 mai. 2023.

CONGER, K. **Driver Charged in Uber's Fatal 2018 Autonomous Car Crash**. THE NEW YORK TIMES. 2020. Disponível em: <https://www.nytimes.com/2020/09/15/technology/uber-autonomous-crash-driver-charged.html>. Acesso em: 17 mai. 2023.

COYLE, D. **Antecipando um monopólio de IA**. Valor Econômico. 2023.

CULBERTSON, D. **US Business-to-Business Labor Market Update – 2023 Q1**. Hiring Lab. 2023. Disponível em: https://www.hiringlab.org/2023/04/25/us-business-to-business-labor-market-update-2023-q1/?utm_source=newsletter&utm_medium=email&utm_campaign=newsletter_axiosmarkets&stream=business. Acesso em: 29 mai. 2023.

DAVENPORT, T; KATYAL, V. **Every Leader's Guide to the Ethics of AI**. MIT Sloan Management Review and SAS. 2019.

DOSI, G. **Technological paradigms and technological trajectories: A suggested interpretation of the determinants and directions of technical change**. Research Policy. Volume 11, Issue 3. 1982. Disponível em: [https://doi.org/10.1016/0048-7333\(82\)90016-6](https://doi.org/10.1016/0048-7333(82)90016-6). Acesso em: 5 jul. 2023.

ELOUNDOU, T *et al.* **GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models**. arxiv. 2023.

ENGLER, A. **Early thoughts on regulating generative AI like ChatGPT**. BROOKINGS. 2023. Disponível em: https://www.brookings.edu/blog/techtank/2023/02/21/early-thoughts-on-regulating-generative-ai-like-chatgpt/?utm_campaign=Governance%20Studies&utm_medium=email&utm_content=247299400&utm_source=hs_email. Acesso em: 18 mai. 2023.

ENGLER, A. **Enrollment algorithms are contributing to the crises of higher education**. BROOKINGS. 2021. Disponível em: <https://www.brookings.edu/research/enrollment-algorithms-are-contributing-to-the-crises-of-higher-education/>. Acesso em: 18 mai. 2023.

ENGLER, A. **The AI Bill of Rights Makes Uneven Progress on Algorithmic Protections**. BROOKINGS and LAWFARE. 2022a. Disponível em: <https://www.lawfareblog.com/ai-bill-rights-makes-uneven-progress-algorithmic->

[protections?utm_campaign=Brookings%20Brief&utm_medium=email&utm_content=229936449&utm_source=hs_email#](https://www.brookings.edu/blog/techtank/2022/05/26/the-eeoc-wants-to-make-ai-hiring-fairer-for-people-with-disabilities/?utm_campaign=Brookings%20Brief&utm_medium=email&utm_content=229936449&utm_source=hs_email#). Acesso em: 18 mai. 2023.

ENGLER, A. **The EEOC wants to make AI hiring fairer for people with disabilities.** BROOKINGS. 2022b. Disponível em: https://www.brookings.edu/blog/techtank/2022/05/26/the-eeoc-wants-to-make-ai-hiring-fairer-for-people-with-disabilities/?utm_campaign=Center%20for%20Technology%20Innovation&utm_source=hs_email&utm_medium=email. Acesso em: 18 mai. 2023.

ENGLER, A. **The EU AI Act will have global impact, but a limited Brussels Effect.** BROOKINGS. 2022c. Disponível em: https://www.brookings.edu/research/the-eu-ai-act-will-have-global-impact-but-a-limited-brussels-effect/?utm_campaign=Brookings%20Brief&utm_medium=email&utm_content=216117153&utm_source=hs_email. Acesso em: 18 mai. 2023.

EWUNETIE, J; HOFHEINZ, P. **What Artificial Intelligence Can Do for Workers and What Skilled Workers Can Do for the World.** Lisbon Council Research sprl. Special Report Issue 01. 2022.

FELTEN, E; RAJ, M; SEAMANS, R. **Occupational Heterogeneity in Exposure to Generative AI.** 2023. Disponível em: <https://doi.org/10.2139/ssrn.4414065>. Acesso em: 5 jul. 2023.

FINQUELIEVICH, S. **No panic. AI will not steal our work overnight.** International conference “Tangible and Intangible Impact of Information and Communication in the Digital Age” Khanty-Mansiysk, Russia. 2019.

FOFONKA, R. **Inteligência artificial: uma poderosa aliada na preservação ambiental.** MIT Sloan Management Review. 2022. Disponível em: https://www.mitsloanreview.com.br/post/inteligencia-artificial-uma-poderosa-aliada-na-preservacao-ambiental?utm_campaign=campanha_forum_sas_data_science_marco_envio_1&utm_medium=email&utm_source=RD+Station. Acesso em: 17 mai. 2023.

FONSEKA, C. **Hold Artificial Intelligence Accountable.** SITN HARVARD. 2017. Disponível em: <https://sitn.hms.harvard.edu/flash/2017/hold-artificial-intelligence-accountable/>. Acesso em: 17 mai. 2023.

FRIEDLER, S; VENKATASUBRAMANIAN, S; ENGLER, A. **How California and other states are tackling AI legislation.** BROOKINGS. 2023. Disponível em: <https://www.brookings.edu/blog/techtank/2023/03/22/how-california-and-other-states-are->

JACKSON, L. **Good morning. Artificial intelligence is changing how people work. That can be scary. But for now, most jobs are still safe.** The new York Times. 2023. Disponível em: https://messaging-custom-newsletters.nytimes.com/template/oakv2?campaign_id=9&emc=edit_nm_20230302&instance_id=86649&nl=the-morning&productCode=NN®i_id=66577200&segment_id=126660&te=1&uri=nyt%3A%2F%2Fnewsletter%2F60f2f70a-aa4f-5724-88b8-64fc9a057da5&user_id=0bc45a3fd8e25ea996c9ced76d21f174. Acesso em: 18 mai. 2023.

JACOBSEN, R. *et al.* **How Chinese Grocers Can Leapfrog the Competition.** BCG. 2022.

KELLEHER, A. **Moore's Law – Now and in the Future.** INTEL. 2022. Disponível em: <https://www.intel.com.br/content/www/br/pt/newsroom/opinion/moore-law-now-and-in-the-future.html>. Acesso em: 17 mai. 2023.

KERRY, C; MELTZER, J; SHEEHAN, M. **CAN DEMOCRACIES COOPERATE WITH CHINA ON AI RESEARCH?.** REBALANCING AI RESEARCH NETWORKS. BROOKINGS. 2023.

KORDA, M. **Could a Chatbot Teach You How to Build a Dirty Bomb?.** OUTRIDER. 2023. Disponível em: <https://outrider.org/nuclear-weapons/articles/could-chatbot-teach-you-how-build-dirty-bomb>. Acesso em: 18 mai. 2023.

KORINEK, A; BALWIT, A. **Aligned with whom?. DIRECT AND SOCIAL GOALS FOR AI SYSTEMS.** BROOKINGS. 2022. Disponível em: <https://doi.org/10.3386/w30017>. Acesso em: 5 jul. 2023.

KORINEK, A; JUELFIS, M. **Preparing for the (non-existent?) future of work.** BROOKINGS. Center on Regulation and Markets Working Paper #3. 2023. Disponível em: <https://doi.org/10.3386/w30172>. Acesso em: 5 jul. 2023.

KORINEK, A; STIGLITZ, J. **ARTIFICIAL INTELLIGENCE, GLOBALIZATION, AND STRATEGIES FOR ECONOMIC DEVELOPMENT.** NATIONAL BUREAU OF ECONOMIC RESEARCH 1050 Massachusetts Avenue Cambridge, MA 02138. 2021. Disponível em: <https://doi.org/10.3386/w28453>. Acesso em: 5 jul. 2023.

LARSEN, B. **The geopolitics of AI and the rise of digital sovereignty.** BROOKINGS. 2022. Disponível em: https://www.brookings.edu/research/the-geopolitics-of-ai-and-the-rise-of-digital-sovereignty/?utm_campaign=Economic%20Studies&utm_medium=email&utm_content=237717337&utm_source=hs_email. Acesso em: 17 mai. 2023.

LAURENT, L. **The AI gold rush will take humanity to dark places.** Bloomberg. 2023. Disponível em: <https://www.businesslive.co.za/bloomberg/news/2023-05-31-the-ai-gold-rush-will-take-humanity-to-dark-places/>. Acesso em: 10 jun. 2023.

LAYOFFS.FYI. **tech companies w/ layoffs.** 2023. Disponível em: <https://layoffs.fyi/>. Acesso em: 29 mai. 2023.

LEE, N; MALAMUD, J. **Opportunities and blind spots in the White House’s blueprint for an AI Bill of Rights.** BROOKINGS. 2022. Disponível em: https://www.brookings.edu/blog/techtank/2022/12/19/opportunities-and-blind-spots-in-the-white-houses-blueprint-for-an-ai-bill-of-rights/?utm_campaign=Governance%20Studies&utm_medium=email&utm_content=238951422&utm_source=hs_email. Acesso em: 17 mai. 2023.

LI, C. **Pelosi’s Visit to Taiwan: Provoking the First AI War in History?.** Chinausfocus. 2022. Disponível em: https://www.chinausfocus.com/foreign-policy/pelosis-visit-to-taiwan-provoking-the-first-ai-war-in-history?utm_campaign=John%20L.%20Thornton%20China%20Center&utm_medium=email&utm_content=224551550&utm_source=hs_email. Acesso em: 17 mai. 2023.

LITTMAN, M *et al.* **Gathering Strength, Gathering Storms: The One Hundred Year Study on Artificial Intelligence (AI100) 2021 Study Panel Report.** Stanford University, Stanford, CA. 2021.

MARTINS, F. **Chat GPT-4: inteligência artificial mente para completar tarefa e gera preocupação.** CNN BRASIL. 2023. Disponível em: <https://www.cnnbrasil.com.br/tecnologia/chat-gpt-4-inteligencia-artificial-mente-para-completar-tarefa-e-gera-preocupacao/>. Acesso em: 18 mai. 2023.

MASLEJ, N. *et al.* **The AI Index 2023 Annual Report.** AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA. 2023.

MAURAN, C. **Samsung bans ChatGPT, AI chatbots after data leak blunder.** Mashable. 2023. Disponível em: <https://mashable.com/article/samsung-chatgpt-leak-leads-to-employee-ban>. Acesso em: 18 mai. 2023.

MAXIME. **What is a Transformer?.** Medium. 2019. Disponível em: <https://medium.com/inside-machine-learning/what-is-a-transformer-d07dd1fbec04>. Acesso em: 18 mai. 2023.

MCKINSEY & COMPANY. **McKinsey Technology Trends Outlook 2022.** 2022

MCKINSEY & COMPANY. **What is generative AI?**. MCKINSEY & COMPANY. 2023. Disponível em: <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai?stcr=ED9D14B2ECF749468C3E4FDF6B16458C&cid=other-eml-alt-mip-mck&hlkid=e591da5dafba4a359fc262342b2ee364&hctky=2092446&hdpid=e4135115-a523-4afa-88b7-778c826b5d42>. Acesso em: 18 mai. 2023.

MCTI. **PORTARIA MCTI Nº 4.979, DE 13 DE JULHO DE 2021**. Ministério da ciência, tecnologia e inovações. 2021.

MINCYT. **Plan Nacional de Inteligencia Artificial**. Ministerio de Ciencia, Tecnología e Innovación. 2019.

MIT TECHNOLOGY REVIEW, 2023. **Building the backbone for innovation, speed, and thriving humanity**. Volume 126, Number 2. 2023.

MOODYS. Credit Outlook 12 September 2022. **Artificial intelligence trial expanding French municipalities' tax base is credit positive**. 2022.

MORDOR INTELLIGENCE. **ARTIFICIAL INTELLIGENCE (AI) MARKET IN AGRICULTURE - GROWTH, TRENDS, COVID-19 IMPACT, AND FORECASTS (2023 - 2028)**. MORDOR INTELLIGENCE. 2021. Disponível em: <https://www.mordorintelligence.com/industry-reports/ai-in-agriculture-market>. Acesso em: 17 mai. 2023.

NELSON, R; WINTER, S. **AN EVOLUTIONARY THEORY OF ECONOMIC CHANGE**. The Belknap press of Harvard University press Cambridge, Massachusetts, and London, England. 1982

NELSON, R; WINTER, S. **Evolutionary Theorizing in Economics**. Journal of Economic Perspective, Vol. 16, NO. 2, Spring. 2002. Disponível em: <https://doi.org/10.1257/0895330027247>. Acesso em: 5 jul. 2023.

OECD. **AI NATIONAL PLAN**. OECD.IA. 2019. Disponível em: <https://oecd.ai/en/dashboards/policy-initiatives/http:%2F%2Faipo.oecd.org%2F2021-data-policyInitiatives-26935>. Acesso em: 17 mai. 2023.

OECD. **BRAZILIAN AI STRATEGY**. OECD.IA. 2023a. Disponível em: <https://oecd.ai/en/dashboards/policy-initiatives/http:%2F%2Faipo.oecd.org%2F2021-data-policyInitiatives-27104>. Acesso em: 17 mai. 2023.

OECD. **Live data**. OECD.IA. 2023b. Disponível em: <https://oecd.ai/en/data?selectedArea=ai-news>. Acesso em: 17 mai. 2023.

PATEL, D; AHMAD, A. **The Inference Cost Of Search Disruption – Large Language Model Cost Analysis**. SemiAnalysis. 2023. Disponível em: <https://www.semianalysis.com/p/the-inference-cost-of-search-disruption>. Acesso em: 18 mai. 2023.

PERSET, K. *et al.* **HARNESSING THE POWER OF AI AND EMERGING TECHNOLOGIES**. BACKGROUND PAPER FOR THE CDEP MINISTERIAL MEETING. OECD DIGITAL ECONOMY PAPERS November 2022 No. 340. 2022.

PETERSON, D; HOFFMAN, S. **GEOPOLITICAL IMPLICATIONS OF AI AND DIGITAL SURVEILLANCE ADOPTION**. BROOKINGS. 2022.

PAVITT, K. **Sectoral patterns of technical change: Towards a taxonomy and a theory**. Science Policy Research Unit. University of Sussex, Brighton BN, UK. 1984. Disponível em: [https://doi.org/10.1016/0048-7333\(84\)90018-0](https://doi.org/10.1016/0048-7333(84)90018-0). Acesso em: 5 jul. 2023.

PRESSE, F. **Presidente da IBM diz que empresa avalia substituir empregos administrativos por inteligência artificial**. G1 Globo. 2023. Disponível em: <https://g1.globo.com/tecnologia/noticia/2023/05/03/presidente-da-ibm-diz-que-empresa-avalia-substituir-empregos-administrativos-por-inteligencia-artificial.ghtml>. Acesso em: 29 mai. 2023.

RANSBOTHAM, S. *et al.* **Artificial Intelligence in Business Gets Real**. MIT Sloan Management Review and The Boston Consulting Group. REPRINT NUMBER 60280. 2018.

RITCHIE, H; ROSER, M. **Age Structure**. Our World In Data. 2019. Disponível em: <https://ourworldindata.org/age-structure>. Acesso em: 25 mai. 2023.

RODRIK, D; STANTCHEVA, S. **Fixing capitalism’s good jobs problem**. Oxford Review of Economic Policy, Volume 37, Number 4, 2021, pp. 824–837. 2021. Disponível em: <https://doi.org/10.1093/oxrep/grab024>. Acesso em: 5 jul. 2023.

SCHUMPETER, J. **CAPITALISMO, SOCIALISMO E DEMOCRACIA**. Editado por George Allen e Unwin Ltd., traduzido por Ruy Jungmann. Rio de Janeiro: Editora Fundo de Cultura. 1961.

SCUDILIO, J. **Como a Inteligência Artificial está transformando os Negócios**. FLAI. 2020. Disponível em: <https://www.flai.com.br/juscudilio/como-a-inteligencia-artificial-esta-transformando-os-negocios/>. Acesso em: 17 mai. 2023.

SEVILLA, G. **More than 100,000 tech workers laid off so far in 2023, but job opportunities are opening up**. Insider Intelligence. 2023. Disponível em: <https://www.insiderintelligence.com/content/more-than-100-000-tech-workers-laid-off-far->

[2023-job-opportunities-opening-up?_ga=2.15547454.967139744.1685330665-](https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-next-frontier-for-ai-in-china-could-add-600-billion-to-its-economy#/)

[1548979214.1683599930](https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-next-frontier-for-ai-in-china-could-add-600-billion-to-its-economy#/). Acesso em: 29 mai. 2023.

SHEN, K. *et al.* **The next frontier for AI in China could add \$600 billion to its economy.**

McKinsey & Company. 2022. Disponível em:

https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-next-frontier-for-ai-in-china-could-add-600-billion-to-its-economy#/. Acesso em: 17 mai. 2023.

SILVER, A. **Deep Blue's cheating move.** en.chessbase. 2015. Disponível em:

<https://en.chessbase.com/post/deep-blue-s-cheating-move>. Acesso em: 17 mai. 2023.

SINGH, J; LUNDEN, I. **OpenAI closes \$300M share sale at \$27B-29B valuation.**

TechCrunch. 2023. Disponível em: [https://techcrunch.com/2023/04/28/openai-funding-](https://techcrunch.com/2023/04/28/openai-funding-valuation-chatgpt/)

[valuation-chatgpt/](https://techcrunch.com/2023/04/28/openai-funding-valuation-chatgpt/). Acesso em: 10 jun. 2023.

STRACK *et al.* **The Future of Jobs in the Era of AI.** BCG and Faethm AI. 2021.

S&P GLOBAL PLATTS. **METALS INSIGHT.** Artificial Intelligence in steel: mills embrace Big Data, robotics. Volume 12, Issue 14. 2017.

THE AI ACT. **The Artificial Intelligence Act.** Artificial Intelligence Act. 2022. Disponível

em: <https://artificialintelligenceact.eu/>. Acesso em: 17 mai. 2023.

THE ECONOMIST GROUP. **Seizing the opportunity: the future of AI in Latin America.** ECONOMIST IMPACT. 2022.

THE WHITE HOUSE. **BLUEPRINT FOR AN AI BILL OF RIGHTS.** MAKING

AUTOMATED SYSTEMS WORK FOR THE AMERICAN PEOPLE. The White House

Office of Science and Technology Policy. 2022.

THEOTO, T. *et al.* **Responsible AI for an Era of Tighter Regulations.** BCG. 2022.

TURING, A. **COMPUTING MACHINERY AND INTELLIGENCE.** Mind, Volume LIX,

Issue 236, Pages 433–460. 1950. Disponível em: <https://doi.org/10.1093/mind/LIX.236.433>.

Acesso em: 5 jul. 2023.

TURNER, M. **Tech giants aren't just cutting thousands of jobs — they're making them**

extinct. INSIDER. 2023. Disponível em: [https://www.businessinsider.com/tech-jobs-arent-](https://www.businessinsider.com/tech-jobs-arent-coming-back-2023-4)

[coming-back-2023-4](https://www.businessinsider.com/tech-jobs-arent-coming-back-2023-4). Acesso em: 18 mai. 2023.

UAE. **100 Practical Applications and Use of Generative AI.** MINISTER OF STATE FOR

ARTIFICIAL INTELLIGENCE, DIGITAL ECONOMY & REMOTE WORK

APPLICATIONS OFFICE. 2023.

UBS SECURITIES AUSTRALIA LTD. **AI and the Rise of Superstar firms.** What does a

disrupted future look like?. 2020

UBS. **US Internet ChatGPT highlights future of consumer AI**. UBS Securities LLC. 2023.

VIEIRA, R. **TEORIA DA FIRMA E INOVAÇÃO: UM ENFOQUE NEO-SCHUMPETERIANO**. Revista Cadernos de Economia. 2010.

VITON, R; LESKA, A; TEIXEIRA, T. **Agtech Innovation Map in Latin America and the Caribbean**. Inter-American Development Bank. 2019.

ZAVERI, P. **6 charts show that tech giants like Meta and Google have still grown like crazy even after layoffs**. Insider. 2023. Disponível em: <https://www.businessinsider.com/impact-of-big-tech-layoffs-amazon-meta-microsoft-google-salesforce-2023-2#-4>. Acesso em: 29 mai. 2023.