**PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO
RELAÇÕES INTERNACIONAIS**

MARIA EDUARDA BEZERRA ALONSO - RA00275663

Os entraves para uma governança internacional da Inteligência Artificial

São Paulo
2024

MARIA EDUARDA BEZERRA ALONSO

Os entraves para uma governança internacional da Inteligência Artificial

Trabalho apresentado à banca examinadora da Pontifícia Universidade Católica de São Paulo, como exigência parcial para obtenção do título de Bacharel em Relações Internacionais, sob a orientação do prof. dr. Arthur Felipe Murta Rocha Soares

São Paulo

2023

AGRADECIMENTOS

Agradeço e dedico este TCC aos meus pais, meus maiores heróis, por todo apoio, amor e persistência para que eu pudesse me trilhar meu próprio caminho. Aos meus avós e a minha tia, pelo carinho, cuidado e confiança durante essa jornada. À minha irmã, por sua força e originalidade. Me inspiro em você tanto quanto se inspira em mim. Aos meus amigos que me acompanharam na jornada, por estarem sempre comigo nos bons e maus momentos. À Wendy, Wandinha e Shiv, pela companhia e amor incondicionais. Ao meu professor orientador, Arthur Murta, por todos os ensinamentos, orientações e por sua infinita paciência. Aos professores com que pude contar durante minha graduação na PUC-SP. Ao meu querido amigo Murilo, por seu imenso apoio, paciência e amparo na árdua tarefa de escrever esse trabalho. À Estela e Gabi, por serem e estarem, sempre.

RESUMO

ALONSO, Maria Eduarda. Os entraves para uma governança internacional da Inteligência Artificial

O presente trabalho tem como objetivo identificar os caminhos apresentados para o desenvolvimento de uma governança da Inteligência Artificial (IA) nos campos teóricos e políticos, a fim de compreender os princípios, estratégias e entraves encontrados pelos principais atores do Sistema Internacional para lidar com a crescente demanda por maior regulação tecnológica. Para tal, a pesquisa passa pelo processo para criação de uma governança da Internet, tecnologia predecessora da IA, a fim de localizar o entendimento teórico acerca das dificuldades na formação de regimes semelhantes. No desenvolvimento do trabalho, a partir do entendimento das diferentes categorias que compõem a tecnologia da IA, são destacados algumas das principais iniciativas estatais e regionais para pôr fim ao vácuo normativo em que se encontra a Inteligência Artificial. Utilizando uma abordagem analítica das Relações Internacionais, investigamos as movimentações globais para a regulação da inteligência artificial, explorando os impactos da sua globalização, para compreender quais caminhos são possíveis para o estabelecimento de sua governança.

Palavras-chave: Inteligência Artificial; Governança; Instituições Internacionais; Internet; Multilateralismo

ABSTRACT

ALONSO, Maria Eduarda. Os entraves para uma governança internacional da Inteligência Artificial

The present work aims to identify the paths presented for the development of governance of Artificial Intelligence (AI) in theoretical and political fields, in order to understand the principles, strategies, and obstacles encountered by key actors in the International System to address the growing demand for greater technological regulation. To this end, the research follows the process for creating governance of the Internet, a predecessor technology of AI, to locate theoretical understanding of the difficulties in forming similar regimes. In the development of the work, based on the understanding of the different categories that compose AI technology, some of the main state and regional initiatives to address the normative vacuum in which Artificial Intelligence finds itself are highlighted. Using an analytical approach from International Relations, we investigate global movements for the regulation of artificial intelligence, exploring the impacts of its globalization to understand which paths are possible for establishing its governance.

Keywords: Artificial Intelligence; Governance; International Institutions; Internet; Multilateralism.

SUMÁRIO

1. INTRODUÇÃO.....	8
2. A GOVERNANÇA DA INTERNET.....	11
3. A INTELIGÊNCIA ARTIFICIAL.....	14
4. OS CAMINHOS PARA UMA GOVERNANÇA DA INTELIGÊNCIA ARTIFICIAL...	17
5. CONCLUSÃO.....	20

1. INTRODUÇÃO

Poucos anos após refletir sobre os aspectos geopolíticos da bomba atômica em *Doutor Fantástico* (1964), o diretor Stanley Kubrick retoma a discussão sobre avanços científicos através do célebre *2001: Uma Odisseia no Espaço* (1968), cuja temática central é a evolução da humanidade e o progresso tecnológico. Dentre os arcos do filme, é apresentado como vilão o computador superinteligente HAL 9000, que, recebendo comandos contraditórios de sempre apresentar informações verdadeiras e esconder da tripulação a verdade sobre a missão para Júpiter, resolve suas problemáticas paradoxais através do assassinato de membros de sua tripulação. HAL é projetado como uma ferramenta de percepção e comunicação onipresente que recebe informações, analisa-as e as projeta rápida e efetivamente. Sua oposição à humanidade por seu comprometimento programado pelo sucesso da missão, ao mesmo tempo pode “imitar” o raciocínio humano, o que resulta ao seu entendimento sobre si como ser consciente e, por tanto, temeroso ao desconhecido. Kubrick utiliza esse elemento para instigar o caos. "*2001: Uma Odisseia no Espaço*", ainda que não seja a primeira obra cinematográfica que trata do confronto entre humanos e máquinas, destaca-se por sua temática transhumanista¹, onde HAL-9000 toma consciência de sua existência e entende-se como ser senciente.

Em 1968, a inteligência artificial (AI), ainda que já teorizada, era presumida como algo futurista. O panorama atual, no entanto, é bastante diferente. A progressão das Tecnologias da Informação e Comunicação (TIC) desencadeou uma nova era, conhecida como a Quarta Revolução Industrial, a qual se destaca das suas predecessoras por sua capacidade de provocar uma ruptura sistêmica que afeta não só as bases da tecnologia, mas também o funcionamento intrínseco da vida política, econômica e social. Dentro dessa evolução, uma série de aspectos notáveis surgiram, abrangendo desde a Internet das Coisas até os serviços em nuvem, incluindo

¹O transumanismo é o posicionamento filosófico de que seres humanos devem ser autorizados a usar a tecnologia para modificar a cognição humana e a função corporal, aprimorando-se para além restrições biológicas. A filosofia comenta também o contexto em que os humanos não mais seriam os únicos seres sencientes, adentrando temáticas de direitos jurídicos à máquinas superinteligentes. Para mais referências sobre o direito de robôs ver L. Young "Are Robots Deserving of Rights?: A critical analysis of how human technological innovation may result in an extension of rights to autonomous cyborg living". Warwick Undergraduate Law Journal. Disponível em: https://warwick.ac.uk/fac/soc/law/aboutus/wulj/article_5.pdf

o Blockchain², Big Data³, redes 5G⁴ e até mesmo a computação quântica⁵. No entanto, há um componente dentro das TIC que tem recebido uma atenção especial devido ao seu alcance e potencial: a Inteligência Artificial. (CARRILO, 2020)

Em uma pesquisa rápida, é possível encontrar diversos artigos sobre como o seu desenvolvimento mudará o mundo, seja em termos de tecnologia, economia, cultura ou mesmo provocando uma mudança estrutural no curso da humanidade. Ela está presente em diversas áreas (WEST & ALLEN, 2018):

A IA não é uma visão futurística, mas sim algo que existe hoje e está sendo integrado e implantado em diversos setores. Isso inclui áreas como finanças, segurança nacional, saúde, justiça criminal, transporte e cidades inteligentes. Existem numerosos exemplos em que a IA já está a causar impacto no mundo e a aumentar as capacidades humanas de formas significativas.

A evolução tecnológica terá também um profundo impacto nas relações internacionais, como apontado pelo Embaixador Eugênio Vargas Garcia (2022), na segurança global e na diplomacia, com muitos prevendo que tais avanços inovadores terão o potencial de reconfigurar o cenário da guerra no século XXI, entre outras ramificações. O nível de inteligência artificial (IA) e a prontidão dos governos e das empresas será significativo na competitividade e, muito provavelmente, se tornará um fator crucial nos investimentos e no crescimento econômico.

²Blockchain é um banco de dados seguro compartilhado através de uma rede de participantes, permitindo que informações atualizadas sejam acessíveis a todos os envolvidos simultaneamente. What is Blockchain? McKinsey & Company, 6 jun. 2024. Disponível em: <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-blockchain>

³O termo "big data" se refere a conjuntos de dados cujo tamanho ultrapassa a capacidade das ferramentas de software de banco de dados convencionais para capturar, armazenar, gerenciar e analisar, e sua definição não se baseia em um limite fixo de terabytes, implicando grande complexibilidade. MANYIKA, J., et al. Big data: The next frontier for innovation, competition, and productivity. McKinsey & Company, mai. 2011. Disponível em: https://www.mckinsey.com/~media/mckinsey/business_functions/mckinsey_digital/our_insights/big_data_the_next_frontier_for_innovation/mgi_big_data_exec_summary.pdf

⁴As redes 5G são a quinta geração da tecnologia sem fio, cuja velocidade é potencializada e a latência reduzida, com maior capacidade de comportar os rápidos avanços tecnológicos da atualidade. What is 5G?. McKinsey & Company, 8 out. 2022. Disponível em: <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-5g>

⁵De tecnologia superior a computadores digitais, a computação quântica é uma nova abordagem ao cálculo que se baseia em princípios da física fundamental para resolver problemas de alta complexidade com maior agilidade. What is quantum computing?. McKinsey & Company, 5 abr. 2024. Disponível em: <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-quantum-computing>

O advento de uma tecnologia tão atrativa como a IA logicamente abre discussões para além de seu aspecto técnico, criando crescente demanda para a produção teórica e consultiva nos mais diversos domínios acadêmicos. Observa-se latente aumento dos estudos sobre os impactos da inteligência artificial na política doméstica, processo que acompanha as movimentações nacionais em busca de entender e regular o uso dessa tecnologia em seus territórios. No entanto, a produção acadêmica a partir de uma visão internacionalista ainda é tímida, incomparável em quantidade em relação as outras áreas próximas nas ciências humanas e políticas (GARCIA, 2022). A lacuna de estudos sobre a política externa e governança internacional representa não somente um chamado para a apropriação da disciplina a um campo em ascensão, mas uma preocupação natural das Relações Internacionais sobre um fenômeno cujos desenvolvimentos possuem capacidade transformadora para a nossa realidade política, econômica e social. Para mais, o caráter insubordinado das IA, tal qual a internet, a fronteiras nacionais, torna-se o estudo da tecnologia natural para a disciplina, por sua capacidade metodológica de compreender fenômenos transnacionais, possuindo, portanto, as ferramentas para analisar a extensão e natureza da IA nos assuntos mundiais, em um momento onde o internacional e o doméstico não são mais estritamente definidos (MARWALA; NDZENDZE, 2023).

Um estudo preliminar da Unesco (2019 *apud*. GARCIA, 2022) ressaltou que a IA não está confinada a uma localização específica, o que torna a sua regulamentação desafiadora em ambos os âmbitos nacional e internacional. Dada a sua natureza transnacional, soluções devem também ser trabalhadas em escala global. Para mais, destaca-se a urgência para que uma governança da IA não se limite apenas à regulamentação, mas que trabalhe também pela prevenção de práticas nocivas e exploração de possibilidades éticas de seu uso.

Esse processo, no entanto, não é simples, encontrando diversos entraves geopolíticos e econômicos para a sua realização. Essa dissertação trabalhará esse fato através da análise das atuais discussões em fóruns nacionais e internacionais em prol de uma regulação da IA, bem como através de uma revisão teórica sobre a temática, sondando recomendações de acadêmicos e consultorias para a construção de sua governança.

Na segunda seção, será retomado o debate sobre a governança da internet, explorando os desencadeamentos para a sua construção e sua atual configuração. Dessa forma, espera-se encontrar margem comparativa para com a situação das IA, tanto para a compreensão de seu atual momento quanto para as possíveis estratégias de para uma governança própria. Posteriormente, o terceiro capítulo será dedicado a definição detalhada dos conceitos de

Inteligência Artificial e suas aplicações, a fim de atestar complexidade da tecnologia e situá-la na realidade, para além de um desenho ficcional caricato. Por fim, através das iniciativas articuladas pela regulação da IA, bem como pelas discussões teóricas acerca do assunto, serão identificadas as propostas para o desenvolvimento de sua governança e seus entraves no quarto capítulo.

2. A GOVERNANÇA DA INTERNET

A Internet é um exemplo de sucesso dos investimentos em pesquisa e desenvolvimento de infraestrutura de informação. Sua origem remonta às pesquisas iniciais sobre comutação de pacotes na década de 1960, quando J.C.R. Licklider, do MIT, propôs a ideia de uma rede global de computadores interconectados para rápido acesso e processamento de dados e programas. Através de investimentos e apoio científica da *Defense Advanced Research Projects Agency* (DARPA)⁶, esse conceito, semelhante a atual internet, foi materializado via ARPANET, a primeira rede operacional a utilizar a comutação de pacotes, mecanismo que comunica rapidamente dados compartimentados em agrupamentos otimizados. A década de 70 apresentou também grandes avanços para a tecnologia, com a criação do *Network Control Protocol* (NCP), que permitia o desenvolvimento de aplicação da ARPANET, e os revolucionários protocolos TCP/IP, que eram respectivamente responsáveis pelo controle de fluxo de informações e pelo endereçamento e encaminhamento dos pacotes. (LEINER et al., 2009)

Durante este período inicial, a tecnologia predecessora a internet era desenvolvida com foco a uso militar. Os anos 1980 marcaram o período de transição difundida da ARPANET para TCP/IP, de modo que a rede foi dividida em uso militar (MILNET) e de pesquisa acadêmica e científica (ARPANET), sendo está expandida para uso em universidades. A comercialização da Internet ganhou força no final dos anos 1980, quando o encorajamento do uso comercial da infraestrutura da Internet foi matriz para o nascimento de redes privadas. Nos anos 1990, a rápida disseminação de navegadores e da tecnologia da *World Wide Web* (WWW) facilitou o acesso à informação global, transformando a Internet em uma plataforma essencial para novos serviços comerciais e para a comunicação global. (LEINER et al., 2009)

A Internet continua a se expandir para suportar novos serviços. A literatura mais crítica

⁶ Criada em fevereiro de 1958 com o acrônimo inicial ARPA, essa organização estadunidense foi uma resposta do presidente Eisenhower aos avanços tecnológicos soviéticos, especialmente no que tange a corrida espacial. Seu objetivo original era manter a soberania tecnológica do país. Ver: <https://www.darpa.mil/about-us/about-darpa>.

aponta como desafio para o futuro o gerenciamento de seu processo contínuo de mudança e evolução. No passado, o compartilhamento de conhecimento entre instituições foi essencial para seu desenvolvimento, como através dos *Request for Comments* (RFCs), documentos que promoveram um ciclo de feedback entre pesquisadores de diferentes localidades. Para mais, Vint Cerf, um dos pioneiros da Internet, criou várias entidades para coordenar atividades internacionais e gerenciar o crescimento conjunto da Internet (LEINER et al., 2009). Para a sua evolução, foi essencial o desenvolvimento de mecanismos e práticas que, compreendendo seu objetivo de interconexão global, tratassem de seu progresso como interesse conjunto, em que houvesse margens para trabalho colaborativo transnacionalizado.

No que se refere a literatura própria à governança da internet, percebem-se divergências teóricas acerca das temáticas que seriam englobadas dentro dessa denominação. Como apontam Van Eeten e Mueller (2012), a literatura acadêmica sobre o campo da governança da Internet frequentemente se concentra quase exclusivamente em instituições internacionais que discutem explicitamente a governança global da Internet, ao passo que o termo “governança da Internet” não costuma ser aplicado aos estudos de muitas atividades as quais moldam e regulam o funcionamento da Internet.

O Grupo de Trabalho das Nações Unidas sobre Governança da Internet definiu governança da Internet como "o desenvolvimento e a aplicação pelos governos, pelo sector privado e pela sociedade civil, nos seus respectivos papéis, de princípios, normas, regras, procedimentos de tomada de decisão e programas partilhados que moldam a evolução e a utilização da Internet." (WGIG, 2005). Os autores supracitados, no entanto, acreditam que essa determinação seja insuficiente, tendo em vista que, ainda que sejam reconhecidos atores não governamentais no texto, a prática e a literatura formal sobre o domínio acabam por focar estritamente em instituições formais governamentais formadas para o assunto, como a Corporação da Internet para Atribuição de Nomes e Números (ICANN), Autoridade para Atribuição de Números da Internet (IANA) e a Força Tarefa de Engenharia de Internet (IETF) (2012).

Os Estados naturalmente se reúnem em torno dessas instituições centralizadas e formalizadas. Surge então a preocupação de que essas instituições sejam erroneamente estudadas como os principais (e únicos) arranjos para a governança da Internet, de modo que a influência dos Estados seja inflacionada em um sistema distribuído em grande escala global majoritariamente em regimes não formalizados. As decisões e ajustes descentralizados realizados por provedores de serviços de internet, entre outras organizações que administram redes e diversas jurisdições, não são considerados como parte do mesmo processo, embora,

frequentemente, esses tenham impactos muito mais significativos na evolução e utilização da Internet do que organismos como o ICANN (2012).

Para entender essa dinâmica, Van Eeten e Mueller dividem o campo da Governança da Internet em quatro grandes áreas, que individualmente podem ser interpretadas como campos distintos: 1) Governança da Internet, único campo reivindica o termo, inclui o trabalho interdisciplinar da economia institucional, ciência política e relações internacionais, estudos de comunicação e informação, sociologia e direito; 2) Política de telecomunicações, que se abrange a pesquisa sobre regulação de telecomunicações orientada para políticas relacionadas à Internet; 3) Economia da segurança da informação, em que se entende que os problemas de segurança da Internet não são simplesmente técnicos, mas são impulsionados pelos incentivos econômicos dos atores e empresas; e 4) Ciber-direito, tratando-se de direitos autorais, privacidade e censura no ciberespaço, sendo também o espaço de debate a assuntos relativos ao ICANN.

O campo que hoje se identifica como governança da Internet formou-se através de um processo histórico específico. Originalmente, o termo referia-se aos problemas institucionais e políticos da coordenação global de nomes e endereços de domínio na Internet, culminando na criação da ICANN em 1998. A necessidade de revisar o conceito de governança da Internet vai além de um problema acadêmico. Governos, ONGs e outros stakeholders são frequentemente levados ao erro ao presumir que governança da Internet se resume ao que acontece nos fóruns supracitados, não levando em considerações atores terceiros que não integram essas redes. Essa concepção errônea é reforçada por grande parte dos trabalhos acadêmicos realizados sob essa denominação. Dessa forma, os autores argumentam que compreensão mais precisa da governança da Internet beneficiaria as comunidades de pesquisa de ambos os lados do rótulo de governança da Internet, fazendo-as considerar o trabalho umas das outras, além de ajudar governos e outros stakeholders a entender como a Internet é realmente governada e o que isso implica para o cumprimento de seus mandatos e objetivos (2012).

Com o advento da internet e o aumento da capacidade de processamento computacional, a acumulação e análise de vastos volumes de dados tornaram-se possíveis, abrindo novas perspectivas para a compreensão humana. Entre essas, destaca-se o desenvolvimento da inteligência artificial: uma tecnologia que consegue criar e resolver problemas complexos e abstratos de maneiras que imitam os processos cognitivos humanos.

3. A INTELIGÊNCIA ARTIFICIAL

Henry A. Kissinger (2018), teórico das Relações Internacionais e ex-Secretário de Estados dos Estados Unidos, argumenta que a IA pode gerar resultados inesperados, alterar processos de pensamento e valores humanos, e alcançar seus objetivos sem conseguir explicar os fundamentos de suas conclusões. O autor alerta que, atualmente, os esforços de governança tendem a priorizar as aplicações da IA em segurança e inteligência, negligenciando a exploração das mudanças que ela já está provocando na condição humana. Similarmente, o físico teórico Stephen Hawking que sucesso na criação da IA pode ser o maior evento na história de nossa civilização numa perspectiva ao mesmo tempo positiva e negativa, pois acompanhada de seus benefícios, perigos como poderosas armas autônomas também viriam a aparecer (WOOLIANS, 2016),

Como demonstrado por Kubrik (1968), a tecnologia e seus avanços já fazem parte do presente e de um futuro imediato, destinados a transformar o ser humano e seu papel em sociedade. Não é a primeira vez na história que o avanço tecnológico exige uma resposta jurídica, mas a IA representa, sob essa perspectiva, um fenômeno especialmente disruptivo (CARRILLO, 2020). Mas afinal, o que é a Inteligência Artificial?

Os autores Shubhendu e Vijay a explicam da seguinte forma:

A IA é um ramo da Ciência da Computação preocupado com o estudo e a criação de sistemas computacionais (...) para criar máquinas que respondem a estímulos de maneira consistente com as respostas tradicionais dos humanos, considerando a capacidade humana de contemplação, julgamento e intenção. (...) Produzidas pela habilidade e trabalho humanos, essas máquinas devem conduzir-se em concordância com a vida, o espírito e a sensibilidade, embora, na realidade, sejam imitações. (SHUBHENDU; VIJAY, 2013, p. 1, tradução nossa)

O termo, portanto, parte da junção conflituosa da ideia de inteligência, frequentemente considerada uma qualidade exclusiva dos seres humanos, com o adjetivo "artificial", que se opõe ao natural. No entanto, encontram-se divergências conceituais que impedem uma definição unificada e amplamente aceita nos âmbitos institucionais ou acadêmicos. As principais discordâncias se encontram no uso do termo inteligência, uma vez que muitos teóricos, como Luc Julia (2019 *apud*. CARRILLO, 2020) não acreditam que as chamadas IAs são ou possam ser de fato inteligentes, no sentido consciente da palavra, mas sim que são ferramentas de “competências extraordinárias” (MANTARÁS, 2019 *apud*. CARILLO, 2020).

Ambas básicas teóricas e práticas da IA podem rastreadas a Alan Turing (1912-1954), matemático considerado o pai da computação. Através do desenvolvimento da máquina que viria a decodificar os até então indecifráveis códigos nazistas durante a 2ª Segunda Guerra Mundial, projeto custeado pelo governo britânico, Turing teve a oportunidade e recursos para expandir seus questionamentos para além do desenvolvimento de computadores previamente programados para execução de tarefas. Seu foco passou à possibilidade de máquinas pensarem e aprenderem por conta própria. Turing discutiu três estratégias principais chegar a esse objetivo: 1) IA por programação, envolve a criação de inteligência artificial por meio de programação manual, onde os sistemas são desenvolvidos com base em regras e algoritmos codificados diretamente pelos programadores; 2) IA por aprendizado de máquina *ab initio*, que se foca no aprendizado automático, permitindo que a máquina adquira conhecimentos a partir de dados brutos, sem depender de regras codificadas manualmente; 3) IA usando lógica, probabilidades, aprendizado e conhecimento de fundo, combinando técnicas como lógica formal, teoria da probabilidade, aprendizado automático e uso de conhecimento pré-existente (MUGGLETON, 2014). Essas três estratégias delineadas por Turing formam a base de diferentes fases da pesquisa em IA que têm sido conduzidas ao longo dos últimos cinquenta anos, cada uma contribuindo de maneira distinta para o desenvolvimento do campo e complementando-se mutuamente na busca pela criação de máquinas pensantes.

Atualmente, acadêmicos e empresas de tecnologia como a IBM⁷ (2023) elencam 3 principais categorias para definir a Inteligência Artificial, com base em suas capacidades. Aqui está um resumo das três categorias de IA baseadas em suas capacidades. A primeira delas é nomeada *Narrow AI*, projetada e treinada para realizar uma tarefa específica ou um conjunto restrito de tarefas, geralmente com eficiência e rapidez superiores às capacidades humanas. No entanto, sua atuação é limitada ao domínio para o qual foi desenvolvida e não pode operar fora desse escopo. A segunda, *General AI*, seria capaz de realizar novas tarefas em diferentes contextos, sem necessidade de re-treinamento humano, utilizando-se de habilidades adquiridas previamente, demonstrando uma flexibilidade e adaptabilidade comparáveis à inteligência humana. Finalmente, a *Super AI*, seria capaz de realizar raciocínios, aprender e tomar decisões, como também possuiriam habilidades cognitivas que superariam as dos seres humanos, como entender sentimentos e experiências humanas, desenvolver emoções, necessidades, crenças e

⁷ International Business Machines Corporation, empresa estadunidense do mercado da informática.

desejos próprios, operando em um nível de cognição e compreensão muito além das capacidades humanas atuais.

Dos três níveis elencados, as tecnologias atualmente disponíveis, como ChatGPT da OpenAI e assistentes virtuais como Alexa da Amazon, compõe o grupo *Narrow AI*, de modo que os outros dois encontram-se estritamente ao campo teórico (*ibid.*, 2023). Dentro deste, são subdivididas classes por formação e funcionalidade para diferentes sistemas de Inteligência Artificial que são contemporaneamente desenvolvidos. As classificações seguintes foram elencadas pelo *AI Act*, minuta oficial da primeira lei referente a Inteligência Artificial da União Europeia (UE), com proposta baseada em riscos no desenvolvimento e aplicação da tecnologia, bem como o seu impacto humano (2023).

- a) *Rule-Based AI*: sistemas que seguem um conjunto de regras predefinidas por especialistas para tomar decisões ou realizar ações em um campo específico;
- b) *Machine Learning*: utilizam algoritmos para aprender a partir de dados, sem necessidade de programação explícita, de modo a melhorar seu desempenho individualmente com base na experiência e repetição;
- c) *Deep Learning*: aprendizado de máquina que envolve redes neurais artificiais para aprender a partir de grandes bases de dados e resolver problemas complexos, como processamento de linguagem natural e reconhecimento de imagem;
- d) *Natural Language Processing*: utilizada em assistentes virtuais e serviços de tradução, permite aos computadores entender, interpretar e manipular a linguagem humana;
- e) *Computer Vision*: permite aos computadores interpretar e entender informações visuais do mundo ao seu redor, como em reconhecimento facial;
- f) *Robotic*: o uso de robôs para realizar tarefas de risco, de alta dificuldade ou repetitivas para humanos;
- g) *Expert System*: imitam a capacidade de tomada de decisão de um especialista humano em um determinado domínio, comum em meios financeiros;
- h) *Cognitive Computing*: semelhante à categoria anterior, imita como o cérebro humano funciona para resolver problemas complexos e;
- i) *Generative AI*: usado para criação de conteúdo, como textos, imagens, vídeos e músicas;

Em constante evolução, é esperado o surgimento de modelos cada vez mais completos e complexos de Inteligência Artificial, com maiores capacidades de provocar mudança na vida humana e social. Concomitantemente, surgem as preocupações acerca de seus usos, sejam civis, militares, políticos ou econômicos. Cada vez mais urge a demanda por legislações que norteiem

esse desenvolvimento desenfreado, como é a *AI Act* da EU (*ibid.*). Torna-se também necessário, portanto, o entendimento das discussões que circundam essa problemática.

4. OS CAMINHOS PARA UMA GOVERNANÇA DA INTELIGÊNCIA ARTIFICIAL

Henry Kissinger (2018), ao se deparar com a problemática do rápido progresso da Inteligência Artificial, sintetiza algumas das principais preocupações de ativistas quanto a falta de restrições deste processo:

A era da internet em que já vivemos prefigura algumas das questões e problemas que a IA só tornará mais agudos. O Iluminismo buscou submeter verdades tradicionais a uma razão humana libertada e analítica. O propósito da internet é ratificar o conhecimento por meio da acumulação e manipulação de dados em constante expansão. A cognição humana perde seu caráter pessoal. Os indivíduos se transformam em dados, e os dados se tornam dominantes (KISSINGER, 2018, tradução nossa)

Esse novo paradigma levanta preocupações sobre como a IA impactará a cognição humana e a organização social. A IA tem o potencial de resolver problemas complexos que anteriormente exigiam a inteligência humana, mas faz isso de uma maneira muitas vezes incompreensível para os próprios seres humanos. Isso cria um cenário em que as decisões tomadas por sistemas de IA podem não apenas ser difíceis de entender, mas também podem alterar fundamentalmente os processos de pensamento e os valores humanos.

Margarita Robles Carrillo (2020) categoriza as preocupações com a IA em dois horizontes principais: as trajetórias de longo prazo e as políticas de curto prazo. No horizonte de longo prazo, há um debate sobre a possibilidade de desenvolver uma inteligência artificial geral para além da atual *Narrow AI*, alcançando as apenas idealizadas *General AI* e *Super AI*. Ou seja, a primeira categoria trata do cuidado para com IAs que não apenas imitem, mas superem a inteligência humana. Essa superinteligência poderia representar um risco existencial para a humanidade (*ibid.*, 2020), como alertado também por figuras como Elon Musk⁸ e Stephen Hawking⁹, por suas capacidades de ultrapassar o controle humano.

⁸ Elon Musk and Others Call for Pause on A.I., Citing ‘Profound Risks to Society’, New York Times, mar. 2023. Disponível em: <https://www.nytimes.com/2023/03/29/technology/ai-artificial-intelligence-musk-risks.html>

⁹ Stephen Hawking warns artificial intelligence could end mankind. BBC News, dez. 2014. Disponível em: <https://www.bbc.com/news/technology-30290540>.

Por outro lado, a política de IA a curto prazo aborda as necessidades técnicas imediatas nas áreas da segurança, padronização conceitual, métricas de desempenho, normas, e políticas de governança. Isso inclui a definição de medidas para garantir que os sistemas de IA sejam desenvolvidos e utilizados de maneira ética, responsável e segura. A falta de regulamentos claros pode resultar em implicações legais, éticas e sociais significativas, que variam desde a privacidade de dados e violações de direitos autorais até a responsabilidade em casos de falhas ou mal uso da tecnologia. (CARRILLO, 2020)

Para mais, outro desafio para uma governança uniforme é maneira como essa tecnologia é distribuída entre os países, de forma bastante desigual, criando disparidades significativas no entendimento e na capacidade de regular a IA. Muitos Estados não possuem o conhecimento tecnológico necessário ou os recursos econômicos para desenvolver ou implementar políticas eficazes de governança para IA. (*ibid.*, 2020)

Isso cria uma lacuna epistêmica e política, onde a tecnologia avança em um ritmo mais rápido do que a capacidade das instituições de acompanhar, avaliar e medir seu impacto. Diante desses desafios, várias propostas e modelos de governança foram sugeridos para regular a IA de forma eficaz, cada uma oferecendo uma estrutura para a governança global da IA, embora com diferentes graus de integração e coordenação.

A ONU, através de seu relatório "*A United Nations system-wide strategic approach and road map for supporting capacity development on artificial intelligence*,"¹⁰ propõe um modelo de governança multissetorial que associa a política de IA aos Objetivos de Desenvolvimento Sustentável (ODS). Essa abordagem visa criar uma estrutura onde diversas partes interessadas, incluindo governos, organizações internacionais, setor privado e sociedade civil, possam colaborar para desenvolver e implementar políticas de IA. (GARCIA, 2022)

Outra proposta relevante é a do Grupo de Alto Nível sobre Cooperação Digital¹¹, que sugere uma governança inclusiva, antecipatória e responsiva. Essa abordagem, de caráter mais

¹⁰ Desenvolvida e aprovada pelo Comitê de Alto Nível para Programas, visa apoiar os países em desenvolvimento no contexto da IA à luz dos Objetivos de Desenvolvimento Sustentável. Disponível em: <https://unsceb.org/united-nations-system-wide-strategic-approach-and-road-map-supporting-capacity-development>

¹¹ Ver: "Roteiro para Inclusão Digital", Organização das Nações Unidas. Disponível em: https://www.un.org/en/content/digital-cooperation-roadmap/?_gl=1*_rjldj0*_ga*NTY2OTc2NzI2LjE3MTgyODgyODc.*_ga_S5EKZKSB78*MTcxODI4ODI4Ny4xLjEuMTcxODI4ODMxNi4zMS4wLjA.*_ga_TK9BQL5X7Z*MTcxODI4ODI4Ny4xLjEuMTcxODI4ODMxNi4wLjAuMA

institucional, defende a utilização dos princípios do Direito Internacional e sugere três modalidades de gestão: a criação de uma agência especializada da ONU, uma organização vinculada, ou um órgão subsidiário da Assembleia Geral das Nações Unidas. (*ibid.*, 2022)

Para além das propostas que envolvem diretamente o sistema ONU, no cenário internacional, a governança da IA se desenvolve principalmente através de fóruns regionais ou inter-regionais, como a União Europeia (UE), a Organização para a Cooperação e Desenvolvimento Económico (OCDE), o Grupo dos Sete (G7), e o Grupo dos Vinte (G20). Nota-se, a partir da configuração dos grupos, o domínio do tema por países predominantemente no norte global. Ao elencar as principais iniciativas nacionais e inter-regionais pela regulação da IA, Embaixador Eugenio Vargas Garcia, percebe e expõe a ausência de países do sul global neste debate. (*ibid.*, 2022).

A exemplo das iniciativas supracitadas, em 2019 a OCDE¹² adotou uma recomendação sobre IA não vinculativa que estabelece princípios e diretrizes importantes para o desenvolvimento e uso responsável da tecnologia. Essa recomendação foi endossada pelo G-20, e refletiu-se na "Visão Comum de *Charlevoix* para o Futuro da Inteligência Artificial"¹³ adotada pelo G7, que busca promover uma IA centrada no ser humano, justa e transparente. (*ibid.*, 2022).

A União Europeia tem sido particularmente ativa na área de governança da IA. O *AI Act*, citado na seção anterior deste trabalho, e o *White Paper on AI*¹⁴ são exemplos de esforços para criar um ambiente regulatório, que protege os direitos dos cidadãos e aborda preocupações éticas ao mesmo tempo que promove a inovação. (*ibid.*, 2022).

Apesar dessas iniciativas, são poucos os Estados que possuem estratégias ou planos de ação abrangentes para a IA. Isso se deve, em grande parte, à falta de capacidade tecnológica e econômica em muitos países, o que limita sua capacidade de desenvolver e implementar políticas eficazes de IA. As atividades das Nações Unidas e outros grupos, embora importantes, muitas vezes são setoriais e fragmentadas, focando em temas específicos sem abordar a regulação da IA de maneira integral e global. A cooperação internacional tende a se manifestar em iniciativas regionais ou protagonizadas por países desenvolvidos, o que pode resultar em uma monopolização do debate e na criação de normas que refletem interesses específicos, em vez de atender às necessidades globais. (CARRILLO, 2020; GARCIA, 2022)

¹² Disponível em: <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449>

¹³ Disponível em: https://www.international.gc.ca/world-monde/assets/pdfs/international_relations_relations_internationales/g7/2018-06-09-artificial-intelligence-artificielle-en.pdf

¹⁴ Disponível em: https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

Para enfrentar esses desafios, os autores urgem (*ibid.*, 2020; 2022) a adoção de uma abordagem proativa e aberta para a governança da IA, baseada no Direito Internacional vigente e nos mecanismos normativos existentes. Isso inclui a criação de estruturas regulatórias que possam evoluir com a tecnologia, garantindo que a IA seja desenvolvida e utilizada de maneira que bem-intencionada ao mesmo tempo em que mitiga os riscos potenciais.

5. CONCLUSÃO

À medida que a internet e o aumento do poder computacional facilitaram a acumulação e análise de dados, surgiram perspectivas inéditas para a compreensão humana. Não é necessário esperar por condições ideais para abordar o problema. Um dos principais problemas para a organização de uma governança da Inteligência Artificial se encontra na dificuldade em se estabelecer um princípio comum norteador para tal.

Atualmente, predominam-se perspectivas institucionalistas liberais no meio acadêmico que realiza o esforço de idealizar um caminho para essa regulação. Sem um ordenamento adequada, governos e empresas privadas podem acelerar o desenvolvimento da IA em busca de vantagens estratégicas, ignorando considerações legais, éticas e de segurança, o que poderia ocasionar ainda a desconfiança entre Estados. Assim, argumenta-se (GARCIA, 2022) que um regime bem projetado e cobrindo questões coerentes poderia ser vantajoso e que normas comuns são essenciais para promover interações ordenadas, beneficiando todas as partes envolvidas.

A partir dessa premissa, as discussões sobre uma governança da IA acabam por gerar um outra problemática: a participação de países do sul global nos principais debates e arranjos. Ainda que haja muito a ser feito na prática, princípios normativos acerca de IA são cada vez mais estabelecidos em fóruns como OCDE, G7, EU e ONU. Iniciativas intergovernamentais são essenciais para coordenar a aplicação de princípios compartilhados e evitar os nacionalismos da IA, prevenindo que cada território tenham suas próprias regras incompatíveis comparativamente. No entanto, esse esforço se concentra nos países desenvolvidos ocidentais, com a notável ausência de grandes jogadores como China e Rússia, e a sub-representação de países em desenvolvimento. Considerando que os regimes de governança criados pelos principais países desenvolvedores de IA afetam o Sul Global, surge o argumento da inclusão desses países no centro do debate (GARCIA, 2022). São vários os motivos listados para tal, como a possibilidade de novas perspectivas que não seriam levantadas pelos hegemons do

assunto, ao mesmo tempo que surgem os contrapontos quanto a participação de um maior número de países nas tomadas de decisão uma vez que, concebendo que uma melhor coordenação ocorre frequentemente entre aqueles com objetivos comuns, um número reduzido de Estados nas votações resultaria na maior facilidade para alcance de consenso (ADAN, 2023).

A prontidão para a IA desempenhará um papel crítico na competitividade econômica global, criando uma divisão entre países preparados e não preparados para essa tecnologia. Atualmente, nenhum país da América Latina, Caribe ou África figura entre os vinte mais bem preparados para a IA, segundo o Índice de Prontidão para Inteligência Artificial Governamental. Essa falta de prontidão pode acentuar a desigualdade global, com países tecnologicamente avançados ganhando ainda mais vantagem sobre aqueles que ainda lutam para adotar e regular a IA. (GARCIA, 2022)

Para mais, muitos debates sobre a governança da IA operam com base em pressupostos equivocados, como a existência de uma ética universal ou de valores homogêneos aplicáveis a todas as sociedades. Pelo contrário, não existe uma ética global que atenda igualmente a todas as culturas e ideologias. Autoras com Margarita Carrillo (2020) e Rachel Adams (2021) argumentam contra o estabelecimento de uma governança da IA baseada em ética por estes mesmos motivos. Ao passo que Carrillo indica mais adequado a uma governança baseado no Direito Internacional, Adams ousa ao encorajar, em uma perspectiva mais crítica, ao ideal de que se deve imaginar um mundo onde a humanidade desista da Inteligência Artificial, ao concluir que seria impossível desvincular seus impactos negativos de suas raízes coloniais, centradas no ocidental.

Ainda, a governança da Inteligência Artificial encontra entraves semelhantes aos da Internet, no que se refere a um debate centrado no Estado e instituições internacionais, muitas vezes ignorando de forma deliberada ou não a participação de poderosos atores privados nas principais tomadas de decisão quanto ao rumo da IA. Para que o debate sobre o papel do Estado na governança da IA avance, é necessário resolver problemas fundamentais. Estados tendem a se agrupar em torno de instituições centralizadas, e estudar essas instituições como os principais arranjos para ambas as governanças pode superestimar a influência dos governos, comparado ao papel que de outros atores dominantes na economia global tais quais as *Big Techs* (MOROZOV, 2018).

"2001: Uma Odisseia no Espaço", de Stanley Kubrick (1968), explora as implicações de tecnologias avançadas e a evolução da inteligência. No filme, o ponto de esclarecimento

principal surge após o ser humano, dominado pela tecnologia, a domina de volta, em um desesperado ato de preservar a sua sobrevivência uma vez unicamente garantida pela máquina. A governança da IA também deve enfrentar os desafios éticos e práticos de sua integração na sociedade global. Assim como o filme coloca a humanidade diante de uma inteligência superior, a IA contemporânea exige que governos e organizações internacionais tomem decisões ponderadas e inclusivas para garantir que os avanços tecnológicos sejam geridos de maneira que beneficiem a humanidade como um todo. Diante de alertas desesperados como o Hawking (WOLLIANS, 2016), que teme o fim da humanidade caso essa questão não seja tratada com a devida seriedade, fundamental que essas discussões sejam abrangentes e inclusivas, incorporando as vozes de todas as nações para construir um futuro mais equitativo e seguro na era da inteligência artificial.

REFERÊNCIAS

- 2001 - Uma Odisseia no Espaço. Direção de Stanley Kubrik. Los Angeles: Metro-Goldwyn-Mayer Studios Inc., 1968. 1 DVD (142 min.).
- ADAMS, Rachel. Can Artificial Intelligence be decolonized? **Interdisciplinary Science Reviews**, v. 46, n. 1-2, p. 176-197, 2021. Disponível em: <https://www.tandfonline.com/doi/citedby/10.1080/03080188.2020.1840225?scroll=top&needAccess=true>. Acesso em: 13 jun. 2024
- ADAN, Sumaya Nur. The Case for Including the Global South in AI Governance Discussions. **Governance of AI**, 20 out. 2023. Disponível em: <https://www.governance.ai/post/the-case-for-including-the-global-south-in-ai-governance-conversations>. Acesso em 13 jun. 2024.
- CARRILLO, Margarita Robles. La gobernanza de la inteligencia artificial: contexto y parámetros generales. Revista Electrónica de Estudios Internacionales, n. 29, jun. 2020. Disponível em: https://digibug.ugr.es/bitstream/handle/10481/63035/39_07_Estudio_Robles_Margarita.pdf?sequence=1&isAllowed=y. Acesso em: 13 jun. 2024.
- GARCIA, Eugenio Vargas. Multilateralism and Artificial Intelligence: What Role for the United Nations?. *In*: TINNNIRELLO, Maurizio. The Global Politics of Artificial Intelligence. 1. ed. Taylor & Francis, 2022. p. 57-84. Disponível em: <https://eugeniovargasgarcia.academia.edu/research#publishedbooks>. Acesso em: 13 jun. 2024.
- KISSINGER, Henry A. How the Enlightenment Ends. **The Atlantic**, jun. 2018. Disponível em: <https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/>. Acesso em: 13 jun. 2024.
- LEINER, Barry M. et al. A brief history of the internet. **ACM SIGCOMM Computer Communication Review**, v. 39, n. 5, p. 22-31, out. 2009. Disponível em: <https://doi.org/10.1145/1629607.1629613>. Acesso em: 13 jun. 2024.
- MANYIKA, James et al. Big data: The next frontier for innovation, competition and productivity. **McKinsey & Company**, mai. 2011. Disponível em: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation>. Acesso em: 13 jun. 2023.
- MARWALA, Tshlidzi; NDZENDZE, Bhaso. Artificial Intelligence and International Relations Theories. Cingapura: Palgrave Macmillan Singapore, 2023.
- MOROZOV, Evgeny. Big Tech: A ascensão dos dados e a morte da política. 1. ed. Ubu Editora, dez. 2018.
- MUGGLETON, Stephen. Alan Turing and the development of Artificial Intelligence. **AI Communications**, v. 27, n. 1, p. 3-10, jan. 2014. Disponível em: <https://content.iospress.com/articles/ai-communications/aic579>. Acesso em 13 jun. 2024/
- PINTO, Tayla. Types of AI. **Artificial Intelligence Act**, 12 abr. 2023. Disponível em: <https://artificialintelligenceact.com/understanding-ai-types-of-ai/>. Acesso em: 13 jun. 2024.
- SHUBHENDU, Shukla; VIJAY, Jaiswal. Applicability of Artificial Intelligence in Different Fields of Life. **International Journal of Scientific Engineering and Research**, v. 1, n. 1, set. 2013. Disponível em: <https://www.ijser.in/archives/v1i1/MDExMzA5MTU=.pdf>. Acesso em: 13 jun. 2024.
- VAN EETEN, Michel JG; MUELLER, Milton. Where is the governance in Internet governance? **New Media & Society**, v. 15, n. 5, p. 720-736, nov. 2012. Disponível em: <https://doi.org/10.1177/1461444812462850>. Acesso em: 13 jun. 2024.
- WEST, Darrell M. What is Artificial Intelligence? **Brookings Institution**, 4 out. 2018. Disponível em: <https://www.brookings.edu/articles/what-is-artificial-intelligence/>. Acesso em: 13 jun. 2024.

WOOLIANS, Robert. Stephen Hawking: AI Could Be Human History's Greatest Disaster - But There is an Alternative. **AI Business**, out. 2016. Disponível em: <https://aibusiness.com/responsible-ai/stephen-hawking-ai-could-be-human-history-s-greatest-disaster-but-there-is-an-alternative#>. Acesso em: 13 jun. 2024.

YOUNG, Lucy. Are Robots Deserving of Rights?: A critical analysis of how human technological innovation may result in an extension of rights to autonomous cyborg living. **Warwick Undergraduate Law Journal**. Disponível em: https://warwick.ac.uk/fac/soc/law/aboutus/wulj/article_5.pdf. Acesso em: 13 jun. 2024.