

## **CONTRIBUIÇÕES AO PROJETO DE LEI 21-20 - MARCO LEGAL DA INTELIGÊNCIA ARTIFICIAL NO BRASIL- uma análise inclusiva cosmoética e democrática**

Da mesma forma que importa perguntar por que regular a inteligência artificial (IA), nos cabe questionar o que é ética e o porquê da ética na sua relação com a IA. Os problemas relacionados à IA impõem um diálogo constante entre o Direito, a Filosofia (Ética) e a Tecnologia, já que estamos tratando de temas com características como a da transversalidade, sendo imprescindível a aproximação de campos científicos não jurídicos, resultando numa espécie de equivalente atual do que outrora, ainda há pouco, foi o direito ambiental (CANTARINI, 2020).

Do que se trata, afinal, é de repensarmos a relação entre as diversas disciplinas e saberes, e de discutirmos a inter e a transdisciplinaridade em novas bases (Lúcia Santaella - Cátedra Oscar Sala), diante da dissolução das fronteiras entre as exatas e as humanidades, a exemplo do que ocorre com o Direito Digital, por meio do desenvolvimento de uma teoria inclusiva e democrática, levando-se em consideração o desenvolvimento de uma Teoria Fundamental do Direito Digital e da Inteligência Artificial, aplicando-se a tais temáticas a Teoria dos Direitos Fundamentais, de forma a propiciar uma adequada proteção aos direitos fundamentais envolvidos em tais searas (“Teoria inclusiva dos direitos fundamentais e direito digital”, Paola Cantarini e Willis S. Guerra Filho, Clube de Autores, 2020; “Teoria Fundamental do Direito

digital: uma análise filosófico-constitucional, Paola Cantarini, Clube de Autores, 2020, “Levando os direitos fundamentais à sério” – Migalhas, edição 05.22).

Vivemos na fase da hiperhistória ou pós-história (Vilém Flusser), na sociedade e economia de dados característica da era da 4ª revolução industrial, da indústria 4.0 ou era do silício, ocorrendo a dependência de nosso bem-estar das tecnologias da informação e comunicação, o que diferencia da fase histórica antecedente, na qual indivíduos apenas se relacionavam com tais tecnologias (Luciano Floridi), sem que estivéssemos ainda totalmente dependentes, daí falar-se em “infomania” (Byung-Chul Han).

A tecnologia e em especial a IA como a mais disruptiva das tecnologias, e a cada dia sendo mais utilizada, produz efeitos em todos os setores sociais, na cultura, no ser humano, em nossas subjetividades, e no conceito de ser humano, e com isso no conceito do que significa continuar sendo humano, diante dos novos hibridismos e agenciamentos que surgem na interação tecnologia-humanos. As tecnologias da informação e comunicação se tornam forças ambientais, antropológicas, sociais e interativas, criando e moldando nossa realidade, modificando a forma como nos relacionamos uns com os outros e com nós mesmos, e a forma como interpretamos o mundo. Surge o pós-humano, transformando a forma como nos relacionamos com o nosso em torno, vivendo em espaços com a abolição da distância, surgindo a paradoxal “simultaneidade da presença e ausência, presença

ausente, ou ausência presente” (SANTAELLA, 2013, p. 06).

Nos cabe indagar: os desafios da IA para a ética e o direito serão resolvidos pela própria tecnologia, entendendo-se suficiente a autorregulação regulada, sem uma base via heteroregulação que combine princípios, compliance, governança e via abordagem da risquificação, mas também levando em consideração a inovação e o incentivo a aplicações de IA de risco baixo a direitos fundamentais? É o que parece ser a preocupação central do AI ACT de 04.2021 da união Europeia, buscando um maior equilíbrio entre a proteção aos direitos fundamentais/valores europeus mas de forma a não obstar a inovação, correspondendo à estratégia europeia para a IA de 04/2018 - “IA para a Europa” (COM/2018/237), buscando tornar-se mais competitivos os países da União Europeia, frente a China e aos EUA.

O “logos”, a linguagem é nossa casa, e esta forma de pensamento reflexivo, criativo, imaginativo e sensível, a possibilidade de juízos críticos que nos diferencia como ser humano de outros animais políticos, traz um tipo de reflexão impossível às máquinas, já que estas se concentrariam em oferecer respostas mais próximas de regras, no sentido de não trazerem a observância do contexto, de particularidades e diferenças específicas e socioculturais de determinada realidade, nem tampouco realizam um juízo crítico ou de ponderação. Mas o mais importante é saber fazer as perguntas corretas e que interessam, tais como, como queremos viver, o que nos é

importante como sociedade? O que significa ser, o que significa ser humano no futuro? Qual o futuro do trabalho na sociedade datificada? Qual regulação de IA pretendemos para auxiliar na concretização da dignidade humana e dos valores democráticos, no sentido de uma proteção adequada a direitos fundamentais, que leve em consideração, pois, não somente seu aspecto individual, mas coletivo e social (multidimensionalidade), bem como uma perspectiva inclusiva, com a participação em Conselhos e órgãos de fiscalização de representantes de grupos vulneráveis e que não exclua as diferenças.

A preocupação com a ética na área da inteligência artificial estaria já com seus dias contados, diante da possível ocorrência da “lavagem ética” e da insuficiência dos princípios éticos? Ocorreria a lavagem ética quando as empresas acabam desvirtuando a atenção acerca da necessidade também de uma regulação jurídica na área da inteligência artificial, ao afirmarem ser suficiente apenas um código de condutas, o que de certa forma não contribuiria para a resolução dos problemas, já que não há a necessária imparcialidade e coercitividade como no caso da heteroregulação, muitas vezes não passando de uma carta de boas intenções.

Diante de tais problemáticas, fala-se no fim da era dos códigos de conduta (Luciano Floridi, “The end of an era: from self-regulation to hard law for the digital industry”). Jess Whittlestone neste sentido aponta para a urgência de se encontrar maneiras de incorporar a ética no desenvolvimento e na aplicação da IA, embora até o

momento o catálogo de princípios éticos elaborados por diversos organismos internacionais e empresas tenha se concentrado em princípios gerais, não informando a solução no caso de conflito entre princípios éticos, afirmando a ineficácia dos princípios éticos gerais (TZACHOR, WHITTLESTONE, SUNDARAM, 2020).

Corroboram tais assertivas o estudo denominado “Inteligência Artificial com Princípios: Consenso de Mapeamento”, elaborado pelo Berkman Klein Center for Internet & Society da Harvard Law School (FJELD et al., 2020), traçando um panorama mundial de princípios éticos da IA, concluindo pela existência de uma grande distância entre teoria e prática na articulação dos conceitos e a sua realização concreta; inexistência de elaboração de princípios orientados para aplicações específicas de IA; divergências quanto a conceitos essenciais como, por exemplo, acerca do que se entende por “justiça”.

É essencial em uma regulamentação da inteligência artificial, fundada em uma construção epistemológica, que seja levado em consideração o conceito de ética digital intercultural, as diversas concepções de dignidade humana e de justiça, olhando-se para as particularidades socioculturais do nosso país, fugindo-se de uma lógica ou viés antropocentrista e eurocêntrico, em atenção, outrossim, às Epistemologia do Sul, considerando-se o sul como categoria epistemológica e não geográfica (Boaventura de Souza Santos).

Cada vez mais há a produção de decisões automatizadas em vários setores sociais, criando-se perfis

comportamentais, prática conhecida como *profiling*, relativa a indivíduos e a grupos sociais, a partir de uma imensa quantidade de dados pessoais, sem o necessário consentimento válido, informado, esclarecido, fracionado para cada aplicação e finalidade específica, sem respeito, pois, aos princípios da necessidade e da finalidade, pois os termos de uso e as políticas de privacidade e de cookies de diversos sites são bastante problemáticos, com informações fragmentadas e em linguagem de difícil compreensão para grande parte da população.

Isto é o design muitas vezes estaria sendo utilizado no sentido de desvio de obrigações legais já existentes no campo da proteção de dados, sendo essencial uma métrica específica voltada em direitos fundamentais embutida no design técnico das aplicações de inteligência artificial. A ética significa em seu sentido grego original “postura”, traduzindo em uma postura em relação à vida, a favor da vida, relaciona-se com a postulação epistemológica em termos de estudo, compreensão e de teoria do direito e da filosofia com fundamento nos valores da “poiesis”, e, pois, da poética ( “Teoria Poética do Direito” – Willis S. Guerra Filho, Paola Cantarini), no sentido de abraçar a criatividade, a sensibilidade, a imaginação.

Apesar de alguns filósofos apontarem, a exemplo de Heidegger, para o fim da filosofia após Hegel, é essencial a recuperação do pensamento reflexivo, crítico, interdisciplinar, zetético, indo além de um pensamento reprodutivo de uma série de informações, obtidas em escala crescente, pois este não se confunde com

compreensão, cognição e reflexão, havendo em certo sentido uma relação antípoda entre informação e comunicação, isto é, quanto mais informação menos comunicação e compreensão, diante da inexistência de tempo e de silêncio para a construção do pensamento próprio e autóctone. É o que observou pioneiramente Vilém Flusser (“Vilém Flusser y la cultura de la imagen. Textos escogidos, «Lengua y realidad», Breno Onetto Muñoz, ed., Valdivia (Chile), Universidad Austral de Chile (UACH), 2016) apontando que as coisas estão desaparecendo dando lugar às informações, bem como Byung-Chul Han (“No-cosas”) ao afirmar que estamos em uma fase de transição, da era das coisas para a era das não coisas.

Antecipar princípios éticos, que levem em consideração também a diferença e diversos conceitos de dignidade humana e de justiça, poderá servir para influenciar o design ético da tecnologia, quando valores são designados no design da tecnologia (“ethics by design”). As regulações europeia, canadense e americana já aprovaram princípios para os desenvolvedores de aplicações de IA com vistas ao estabelecimento de “frameworks” de “responsability-by-design”, “privacy-by-design” e “security-by design”.

Verifica-se, pois que é essencial a construção de um sistema de proteção proativo, abrangente e sistemicamente seguro, uma proteção sistêmica, o que envolveria a proteção desde a concepção tecnológica (“protection by design”), por meio da criação de

arquiteturas de decisão adequadas à proteção com o auxílio da concepção e de ferramentas tecnológicas, como forma de se implementar a segurança (“security by design”), falando-se em transparência do design tecnológico (projeto técnico) e dos algoritmos de IA, e não apenas na coleta e tratamento de dados pessoais.

Passa-se do paradigma da autodeterminação informativa para uma arquitetura de gerenciamento de riscos, sendo tal modificação atrelada à abordagem via risquificação, como pode-se observar da nova regulamentação da EU, o “AI Act”, na linha de outros documentos internacionais como o “White Paper on AI”, trazendo diversos níveis de risco quanto a aplicações de IA, elevado, moderado, fraco e intolerável. Um dos pontos a se refletir é se um patamar estabelecido a priori e de forma fixa quanto aos diversos níveis de risco, e não uma abordagem mais flexível, no sentido, de se confirmar o risco diante do caso concreto, seria a melhor abordagem.

Devem também ser observados como possibilidades algumas propostas presentes na Resolução com recomendações sobre regras de Direito Civil e Robótica de 2017 e do Parlamento europeu e na proposta do Parlamento Europeu, de 2015 (2015/2103/INL), embora anteriores, no sentido de um registro obrigatório para robôs, seguros obrigatórios e criação de fundos de compensação, com a estipulação de um patrimônio mínimo autônomo, embora haja algumas críticas no sentido de ser um entrave excessivamente oneroso ao desenvolvimento tecnológico para pequenas empresas,



podendo propiciar o monopólio de mercado. Contudo tal aspecto negativo pudesse ser superado com patamares proporcionais ao faturamento da empresa em questão, de forma proporcional portanto. Referia proposta considera os responsáveis pela reparação de acordo com o nível efetivo de instruções dadas aos robôs e o nível da sua autonomia, na figura denominada de professor, ou seja, do programador. A regra seria da responsabilidade subjetiva, sendo a responsabilidade objetiva de forma excepcional e em situações de acordo com uma lista taxativa, surgindo interpretações na doutrina destoantes do aspecto de taxatividade, contudo.

Tal proposta visaria propiciar uma maior chance de indenização das vítimas de algum dano ocasionado por aplicações de IA, já que nem sempre é fácil se determinar quem seria os responsáveis por tal reparação, como vem apontando a doutrina, a denominada IRRESPONSABILIDADE DISTRIBUÍDA ou o ‘problema de muitas mãos’, pela dificuldade ou falta de identificação do nexos causal entre os diferentes atuantes envolvidos no processo, elemento indispensável da responsabilidade civil. Isso ocorrerá principalmente quando o dano ocorrer dentro de um complexo sistema sociotécnico, no qual não será óbvia a responsabilidade da IA em si, nem de uma pessoa física ou jurídica. Neste sentido iniciativas como do Alan Turing e Oxford Internet Institute/UNESCO falam em uma “responsabilidade ‘compartilhada’ ou ‘distribuída entre designers de robôs, engenheiros, programadores, fabricantes, investidores, vendedores e usuários. Nenhum

desses agentes pode ser indicado como a última fonte de ação. Ao mesmo tempo, esta solução tende a diluir completamente a noção de responsabilidade: se todos tiverem uma parte na responsabilidade total, ninguém será completamente responsável. [...] Os robôs podem ser usados para fins destinados por seus designers, mas também para outros fins, seu 'comportamento' pode ser 'pirateado' ou 'reprogramado' por seus usuários finais. <http://unesdoc.unesco.org/images/0025/002539/253952E.pdf>. (diversas responsabilidades).

Ao invés de pensarmos acerca da essência da técnica objeto de reflexões por parte de Heidegger, em especial ao seu texto de 1949 "A questão da técnica", não em um sentido apenas distópico, como fazendo parte dos dispositivos do biopoder ou do capitalismo de dados, mas, no sentido de refletir acerca da tecnodiversidade e da cosmotécnica, reconhecendo a existência de um pluralismo tecnológico e ontológico, evitando-se a simples oposição dualista entre natureza e técnica, como aponta Yuk Hui, sugerindo que seja repensada a descolonização a partir da perspectiva da tecnopolítica.

Tal postulação reconhece e parte da insuficiência de uma visão eurocêntrica e antropocêntrica para se pensar a relação técnica-humanos, tal como se verifica a exemplo das próprias Declarações Universais dos Direitos Humanos construídas sob uma ótica hegemônica ocidental, não sendo levadas em consideração geralmente as construções das Epistemologias do Sul. Há uma crise quanto aos direitos humanos, como se pode observar do

retrocesso que vem ocorrendo em termos de conquista de direitos e preocupações humanitárias.

Daí se falar da necessidade de uma nova compreensão e da reinvenção dos Direitos Humanos, no sentido de integrar a diversidade cultural e as diversas concepções de justiça e de dignidade humana, a exemplo, das noções de “dharma” Hindu, de “umma” islâmica, de “pachamama” ou o “buen vivir” dos povos indígenas da América Latina, do “ubuntu africano”, do “Sumak Kawsay”, ou o “Sumak Qamanã”, trazendo o respeito aos direitos da natureza, passando do foco dos deveres ao foco aos direitos, e para uma nova concepção de comunidade, a exemplo da Constituição do Equador de 2008, como constitucionalismo transformador.

Tais propostas refletem e são fundamentadas no respeito à diferença, no respeito pela igualdade na diferença, por meio de um processo político participativo, na linha do que se denomina de “constitucionalismo transformador”, trazendo a possibilidade de recuperação da cidadania anestesiada ou passiva, transformada em uma cidadania ativa, como exercício de direitos humanos.

Portanto, ao invés de apenas pensarmos em uma “human-centered IA”, devemos abranger a concepção da tecnologia a favor da vida na Terra de futuras gerações, não apenas, pois, “human-centered IA”, mas “planet centered” ou “life-centered”, em uma visão cosmoética, levando-se em consideração os impactos ao meio ambiente e a crise climática.

Perguntas essenciais que devemos nos fazer por exemplo apontam para reflexões tais como: “Será que preciso de IA para determinada aplicação, considerando os riscos associados, alto grau e energia envolvido e custos? Será que seria possível resolver meu problema específico de outra forma? Algum efeito negativo poderia ser de fato mitigado com aplicação da prevenção e instrumentos adequados de compliance? O quanto de eficiência precisaria embutir em tal aplicação, em termos de custos ambientais envolvidos, como níveis de liberação de carbono e agressão ao meio ambiente?

Trata-se de uma renovação do pensamento jurídico à luz de uma Teoria (Fundamental) do Direito digital e da inteligência artificial, a fim de se possibilitar um maior respeito aos Direitos Fundamentais, voltando-se a uma visão dinâmica do ordenamento jurídico, a partir de uma consideração contextualizada, caso a caso, assegurando-se um procedimento isento, de modo a alcançar decisões aptas a equalizar todos os interesses e/ou valores em conflito, trazendo harmonia e uma solução segura e justa, ante as múltiplas possibilidades de solução. Neste sentido a importância do princípio da proporcionalidade e da ponderação no caso de colisões de normas de direitos fundamentais, diante de conflitos nas áreas de proteção de dados e diante de aplicações de inteligência artificial.

Tal fase e análise estará obrigatoriamente presente dentro da metodologia de um Relatório de Impacto de Direitos Humanos e Fundamentais de aplicações de IA,

assim como se faz presente no Relatório de Impacto de proteção de Dados e na Avaliação do Legítimo Interesse, sendo essencial a elaboração prévia de tais documentos em casos de alto ou moderado grau de riscos à direitos fundamentais e liberdades fundamentais, a ser elaborado por uma equipe interdisciplinar e independente.

Quando se fala pois em questões éticas relacionadas a IA, de uma filosofia da IA, tais temáticas não se limitam ao estabelecimento de códigos éticos ou de evitar vieses de aplicações específicas de IA, mas de repensarmos as bases epistemológicas para a construção do conhecimento científico em tais searas em novas bases, e com fulcro nos valores do “homo poietico” no sentido de uma filosofia, ética, e de um direito libertos do binômio aprisionador sujeito-objeto, mas comprometidos com o múltiplo e o acategórico, no sentido de libertar a diferença, que é o elemento essencial quando se fala em recuperação de diversas características essencialmente humanas, e de modo a reequilibrar a relação humano-algoritmos. Uma leitura e compreensão poéticas, não dialéticas, que levem em conta o não dito, o resto, a heterotopia, superando-se dogmas como o da neutralidade e objetividade da tecnologia, e focando no aspecto da diferença (Tese de doutorado em Filosofia, Paola Cantarini, PUCSP, 2021, “O teatro Filosófico de Foucault e o Direito”).

Tal proposta hermenêutica visa alcançar a perspectiva poética, no sentido de não linear, não bidimensional, trazendo uma leitura não polarizada, mas holística e inclusiva, e a favor de se repensar as

ambivalências e contradições, voltando-se para uma compreensão que passa pelo pensamento filosófico polifônico, do múltiplo, como uma pragmática do múltiplo, um pensamento plural, aproximando-se do que Luciano Floridi (The Logic of Information: A Theory of Philosophy as Conceptual Design) afirma acerca da necessidade de uma lógica de design como uma lógica conceptual que precisa ser desenvolvida para sustentar a base da informação em IA, com base nos valores do "homo poietico". Do que se trata é de uma mudança de uma compreensão do conhecimento representacionalista (mimético) para um construcionista (poiético), da mimesis à poiesis, numa conhecimento representacionalista (mimético) para um construcionista (poiético), da mimesis à poiesis, numa interpretação poiética dos nossos conhecimentos, desenvolvendo uma lógica de "fazer", de design dos artefatos semânticos para os quais nós somos epistemicamente responsáveis.

Chega-se assim a uma filosofia como design conceitual, com fundamentos epistemológicos e metodológicos para um Design ético em IA envolvendo a crítica, uma epistemologia poiética (construtiva) em vez de mimetizada (representativa), apta a propor uma ética da IA relacionada com os valores construcionistas do homo poieticus. Do que se trata então é de pensarmos nas questões de quais os fundamentos e bases epistemológicas e hermenêuticas para a temática da governança de IA e da relação e interação humano-algoritmo, por um lado, respeitando as diferenças, numa

perspectiva multicultural e, por outro lado, de como estabelecer os fundamentos e marcos teóricos para a regulamentação harmônica das tecnologias digitais, mas levando as especificidades brasileiras em consideração? Desta forma, procuramos observar e compreender o objeto de estudo em questão de outra perspectiva, de outro ponto de vista, mas, sobretudo, através de uma visão não polarizada, não dualista e não representativa, mas sim holística e inclusiva, e portanto, democrática, a fim de repensar ambivalências e contradições, assumindo estas e não as ocultando em uma idealização falseadora.

Não há muitas propostas científicas, multidisciplinares e brasileiras voltadas à análise e formulação de frameworks relacionados ao design ético e governança da IA, de modo a estabelecer as bases para a criação de um sistema de proteção aos direitos humanos (DH) e fundamentais (DF) de todas as parcelas da população, que seja proativo, abrangente inclusivo e sistemicamente seguro (proteção sistêmica), envolvendo conceitos como os de Protection by Design, Security by Design, Ethics by Design, Ética Digital Intercultural, tecnodiversidade, cosmoética.

Além da própria filosofia em certo momento desvincular-se da necessidade da busca da verdade como fim último, mesmo porque esta se daria sempre de forma relativa, além do seu caráter de aporia, próprio da filosofia, agora vemos a necessidade também da matemática, e das ciências que se baseiam em tal conhecimento, aceitarem as contradições, as antinomias, próprias de um discurso

auto-referencial, como expôs George Spencer-Brown, abandonando paradigmas já superados como o da separação entre sujeito e objeto do conhecimento, substituindo tal separação por uma unidade, ao invés de “ou”, “e”.

Daí se considerar a importância de teorias transclássicas com foco na abordagem holística e não reducionista, típica das ciências modernas, como a cibernética, a semiótica, a teoria geral de sistemas, as teorias gerais da informação e da comunicação e a cibernética de segunda ordem, tal como proposta por H. Von Foerster, ao descrever sistemas cibernéticos dotados de IA que se autorregulam. Ou seja, na base do conhecimento acerca da IA teríamos uma disciplina transclássica, pós-moderna, fugindo-se do antropocentrismo, da oposição sujeito e objeto e olhando para a diferença e o outro.

Tem-se, pois, por superada, outrossim, a compreensão de uma abordagem do conhecimento apenas compreendendo as contribuições da sociedade ocidental e de uma perspectiva eurocêntrica, como ao se afirmar, por exemplo, que no Oriente não se teria uma filosofia própria sendo esta apenas ocidental, já que a cientificidade necessária estaria atrelada à ideia de uma teoria inclusiva e de interdisciplinaridade, o que demandaria a análise e consideração de um maior número possível de abordagens e perspectivas, de forma democrática. Os direitos humanos, por exemplo, não podem mais ser vistos sob uma única ótica, universalista,



iguais para todo o gênero humano, em uma perspectiva etnocêntrica, ocidental, mas levando-se em consideração as diversas culturas e gêneros, havendo diversas concepções, portanto, de direitos humanos, já que há uma diversidade cultural e social (comunitaristas e multiculturalistas).

Em sentido complementar, os direitos fundamentais, no plano interno voltam-se para uma natureza multidimensional, reconhecendo-se seu aspecto individual, coletivo e social, característica que fica clara ao pensarmos em um vazamento de dados como equivalente a um dano ambiental, causando danos não apenas individuais, já que relacionados à cidadania e à igualdade material dos tutelados. Daí se falar em poluição de dados (BEN-SHAHAR, Omri. *Data Pollution*, p. 133 e ss.), espécie de “direito ambiental da proteção de dados pessoais”. Em sentido complementar Gunther Teubner, traz a advertência de que não basta uma perspectiva individualista na esfera digital, devendo ser buscada sua dimensão coletivo-institucional (TEUBNER, Gunther. 2017, p. 485–510).

Como bem apontam alguns estudos na área de inteligência artificial, que analisaram conjuntos de propostas de codificações e regulações haveria uma ausência de propostas não eurocêntricas, bem como contradições e não compatibilidade quanto ao conceito de justiça, por exemplo, ou de dignidade humana (“Inteligência Artificial com Princípios: Consenso de Mapeamento” - Berkman Klein Center for Internet &

Society da Harvard Law School, FJELD et al., 2020). Também poderíamos apontar outras fragilidades encontradas em algumas propostas de guidelines para IA, tais como se verifica na elaboração da Comissão Europeia, pois de 52 especialistas envolvidos em sua elaboração 23 eram representantes e empregados de grandes empresas, e apenas 4 dos especialistas possuíam conhecimentos em ética, e nenhum em proteção de dados, faltando pois o requisito da representatividade adequada (subrepresentação), e democrática, pois deverá compreender todos os grupos da sociedade.

Outra fragilidade que poderia ser repensada é a de se pensar em direitos fundamentais, sem observância do Constitucionalismo digital, da nova hermenêutica constitucional, do pós-positivismo, adotando-se parâmetros que já não mais fazem sentido, voltados a uma proteção não sistêmica, não proativa, e sem levar em conta direitos coletivos e sociais.

Postula-se, pois por levar-se em consideração nas regulamentações da IA, trazer sempre presente a necessidade de estudos por meio de uma análise multidisciplinar, multidimensional, intercultural já que trata de questões com características polifacetadas, adotando-se uma nova visão hermenêutica e epistemológica, visando à construção de pilares essenciais para o design ético-técnico da IA voltando-se para o "Human and fundamental rights by design", "beneficial AI", "AI for good" e "HumanCentered AI", em uma perspectiva sustentável e não antropocêntrica, mas levando-se em consideração o

meio-ambiente e os outros seres vivos que já são considerados como actantes como bem aponta Bruno Latour.

É essencial pois, uma metodologia própria para a realização de Relatório de Impacto de Inteligência Artificial com base na violação de Direitos Humanos e Direitos Fundamentais, fundamental para uma IA de confiança, levando-se em consideração conceitos vinculados ao Sul Global, tais como propõe Boaventura de Souza Santos quando menciona as Epistemologias do Sul, portanto, de outras concepções não eurocêntricas de dignidade humana, justiça e direitos humanos, de forma a propiciar uma perspectiva inclusiva e democrática, por contribuírem sobretudo para uma visão não antropocêntrica, a fim de fortalecer o mercado nacional de IA e novas tecnologias. Portanto, é essencial olhar para outras perspectivas em termos de “Epistemologias do “Sul”, compreendendo os conceitos de ética digital intercultural, tecnodiversidade e cosmoética, no sentido de superação do monoculturalismo, reconhecendo-se a importância do elemento diferença, da diversidade, não representativa, mas positiva, da diferença.

Busca-se a análise do conceito de justiça algorítmica, ao se propor repensar a questão da técnica, e da essência da técnica com base em valores como da cosmoética, tecnodiversidade, como expõe Yuk Hui (“Technodiversity”, p. 154); voltadas ao empoderamento do ser humano, verificando-se tais questões em outras bases, fora do dualismo e de oposições inconciliáveis, em

uma perspectiva de desenvolvimento da IA inclusiva, democrática e sustentável, além de uma visão antropomórfica, o que não compromete a visão de uma IA voltada ao ser humano, o que estaria compreendida em uma abordagem mais ampla, no sentido de uma proposta sustentável. A questão com o meio ambiente e a sustentabilidade fazem, portanto, parte das discussões éticas que devem ser levadas em consideração. Alguns estudos apontam para o gasto ambiental e energético relacionado a aplicações de IA (Jun 2019, “Energy and Policy Considerations for Deep Learning in NLP, Emma Strubell, Ananya Ganesh, Andrew McCallum), como o estudo de pesquisadores da Universidade de Massachusetts, apontando altos níveis de emissão de dióxido de carbono.

No Brasil, ao lado da Estratégia Brasileira de IA no Brasil, Instituída pela Portaria MCTI nº 4.617, de 6 de abril de 2021, apesar de algumas falhas e omissões, imprecisões técnicas, ausência de obrigações substantivas e processuais, ausência de parâmetros mínimos de procedimentalização e previsão de instrumentos de governança algorítmica, em especial se comparamos com as regulamentações internacionais. Isto porque a autorregulação e mesmo a autorregulação regulada pelas empresas, não seria eficaz nem tampouco contribuiria para o aspecto da confiança, já que muitas vezes tal iniciativa colide com a busca de fins públicos e com a proteção de direitos fundamentais e humanos, voltando-se primordialmente para os valores de mercado, não sendo

iniciativas muitas vezes pautadas na transparência e imparcialidade. Em muitos casos há aqui uma concepção proprietária dos diretos envolvidos, a busca da inovação e de valores econômicos acima de outros valores democráticos, envolvendo a elaboração de conteúdo unilateral e seletivo em termos de interesses, na linha de uma análise econômica do Direito, voltada para eficiência do mercado.

Neste sentido um documento imprescindível e que deverá estar consagrado dentro do texto legislativo de forma a tornar obrigatória sua elaboração prévia, em atenção ao princípio da prevenção, em casos de alto risco e de risco moderado de aplicações de IA, é o Relatório de Impacto algorítmico, evitando-se interpretações dúbias como ocorre atualmente com a redação da LGPD – Lei Geral de Proteção de Dados, por não ter tornado clara tal questão, no caso do Relatório de Impacto de proteção de dados, dividindo a doutrina em entendimentos diversos no sentido de ser obrigatório ou não, qual seu procedimento de elaboração e a partir de qual momento deverá ser elaborado, se apenas quando da requisição da ANPD, ou em uma interpretação literal e não funcional, não voltada para a máxima efetividade e proteção a direitos fundamentais. Uma interpretação literal e restritiva iria ferir toda a lógica de elaboração de tais documentos, pautada na ideia de precaução, de mitigação de riscos, no início do desenvolvimento da tecnologia, já no seu design, não passando pelo crivo de uma análise funcional e sistemática da legislação. Tal interpretação na área de

proteção de dados é contrária ao entendimento do GDPR – Regulamento Geral de proteção e Dados da União Europeia, (RGPD) (UE) 2016/679, o qual foi o principal marco teórico orientador da LGPD, assim como o entendimento e diversas ANPD de diversos países, e orientações de órgãos consultivos como o WP29 - The Article 29 Working Party, e o EDPB - European Data Protection Board.

O princípio da prevenção encontra fundamento na proposta de ética de Hans Jonas, amparada na noção de responsabilidade e de prudência, levando a uma “Eco-responsible AI”, componentes essenciais da ética para o futuro envolvendo, pois, a ética para nossa sociedade pós-moderna, sendo insuficientes os sistemas de ética tradicionais, voltados a uma perspectiva individualista e antropocêntrica apenas.

Há, pois, pelo PL 21-20 uma expressão muito vaga e genérica relacionada à elaboração do relatório de impacto, condicionando sua exigência “a justificção de sua necessidade”, sem maiores comentários ou especificações, contribuindo para a insegurança jurídica e para uma proteção de nível fraco quanto aos direitos fundamentais envolvidos. Deverá prever, pois, de forma obrigatória (alto risco e risco moderado, a depender da aplicação no caso concreto, pois podem ser diversas as consequências e graus de risco de uma aplicação de reconhecimento facial de um ser humano e de um animal) a elaboração de Avaliações de Impacto Regulatório/Relatórios de Impacto a Direitos

Humanos/RIDR e DPIA/RIPDPs), incorporando a noção de efeitos discriminatórios dentro da metodologia de produção dessas avaliações, possibilitando assim a participação de diversas parcelas da população de forma semelhante a um escrutínio público quando envolverem entidades independentes e especialistas para a tarefa de revisão/auditoria de tal aplicação. Deverá possibilitar ainda a verificação de se há uma etapa de *check and balance* dentro da arquitetura tecnológica, compreendendo um teste de proporcionalidade, como no caso do legítimo interesse previsto na LGPD e no GDPR.

Um modelo interessante de metodologia para a elaboração da Avaliação de impacto de IA foi publicado recentemente, denominado “Platform for the information society” ([ecp.nl/wp-content/2019](https://ecp.nl/wp-content/2019)), com destaque para a quarta etapa a análise dos aspectos éticos e legais da aplicação, a etapa 5, acerca do grau de transparência proporcional ao grau do impacto e possível dano a direitos fundamentais, aliado aos critérios da confiança e da segurança, e a etapa 6 trazendo o teste da proporcionalidade, além da etapa 8 afirmando a necessidade da avaliação ser constante e periódica, ou seja, durante todo o ciclo de vida da utilização da tecnologia, como ocorre já, ou deveria ocorrer, com a proteção de dados. Outro modelo elaborado pelo Canadá pode ser obtido no site <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html>.

Deve ainda ser exigido a realização de testes com métodos dinâmicos e testes pré-lançamento, com ênfase na abordagem conhecida como *sandbox approach*, a qual, apesar de prevista no artigo 7º, VII do PL21/20, não possui maiores proceduralizações ou especificações, nem tampouco foi tratada como uma etapa obrigatória de compliance, ou então outra forma de testes, em casos de alto nível de risco ou risco moderado. Seria interessante talvez um incentivo em termos de redução de possíveis penalidades, no caso de tal elaboração ou incentivos tributários, e/ou institucionais. Devem ser ainda levado em consideração o contexto sociocultural brasileiro, onde por exemplo um caso de “bias” envolvendo discriminação ou racismo digital poderá ter maiores repercussões em razão do racismo estrutural (*design AI* para contextos locais).

Para parte da doutrina deveria ser prevista a proibição da utilização de sistemas de “caixa preta” que não possam ser sujeitos a parâmetros significativos de *accountability* e transparência, e proibida sua utilização em sistemas em IA de ALTO RISCO. Deverá estar prevista a revisão dos resultados de decisões automatizadas por meio da análise do processo decisório utilizado pelo algoritmo, ainda mais quando implicar em violações graves a direitos e liberdades fundamentais, e propiciar um procedimento de *due dilligence* para avaliar os aspectos discriminatórios em caso de utilização de dados sensíveis, ou de pessoas de grupos vulneráveis ou sujeita a uma maior vulnerabilidade (idosos, crianças, adolescentes, enfermos). Seria importante em tal procedimento buscar-



se a identificação de potenciais resultados discriminatórios, bem como a realização de ações efetivas para prevenir e mitigar a discriminação/outros riscos, e verificar o grau de transparência no sentido de esforços para identificar, prevenir e mitigar a discriminação.

Outra questão em aberto e já problemática na LGPD e que não é trabalhada pelo PL em questão é a obrigatoriedade da revisão humana de aplicações de IA que proporcionem ou possuam um potencial de risco alto ou moderado a direitos e liberdades fundamentais, seguindo-se a ótica do GDPR, art. 22.3 (“direito de obter intervenção humana, manifestar seu ponto de vista e contestar a decisão”), senão vejamos. Consoante razões do veto presidencial da expressão “pessoa natural”, contrariando o GDPR e os princípios da OCDE, não há mais tal obrigatoriedade, independente de uma análise de risco.<sup>1</sup> Entre os argumentos presentes na justificativa do veto presidencial encontram-se razões utilitárias e econômicas, as quais não seriam as mais apropriadas quando se fala na proteção de direitos fundamentais ou

---

<sup>1</sup> Antes do veto presidencial, a Lei nº. 13.853, de 2019, que aprovou a Medida Provisória 869/2018, alterando a Lei Geral de Proteção de Dados (LGPD) reincorporava à LGPD a obrigatoriedade de revisão de decisões automatizadas por pessoa natural, consoante sua redação original, nos termos do art. 20, e p. 3º, mas foi objeto de veto presidencial.

Importa destacar os principais fundamentos a favor da revisão humana, nos termos do parecer do Relator na Câmara Federal: “(i) a retirada da pessoa humana enfraqueceria o exercício dos direitos humanos, de cidadania e do consumidor previstos no art. 2º, VI e VII da LGPD; (ii) a interação de pessoas com deficiência de julgamento ou falta experiência com controladores seria dificultada, pois a inexistência de contato com revisores humanos poderia levar a práticas abusivas; (iii) que os algoritmos que processam os dados são baseados em cálculos probabilísticos e estatísticas e que, por não englobarem o universo dos titulares e seus comportamentos, poderiam levar a erros e desvios padrões, já que se baseiam apenas em amostras e intervalos de confiança, além de estarem sujeitos a incorreções próprias do desenvolvimento tecnológico; (iv) que a retirada vai de encontro ao que prevê a (colocar o nome completo do ato normativo e depois a Regulamentação Geral de Proteção de Dados – GDPR europeia (equivalente à LGPD brasileira) em seu art. 22, o que poderia dificultar a integração comercial e a geração de oportunidades e investimentos” (...).

direitos humanos, já que estes não podem estar sujeitos exclusivamente a juízos ou critérios de custo-benefício, como já bem demonstrou já a Constituição de Weimar, ao reconhecer que a tutela dos interesses econômicos somente é legítima enquanto ela esteja vinculada aos direitos da pessoa. Ainda mais se tem o sentido de com isso se beneficiar as micro-empresas ou start-ups não trazendo um ônus significativo e reduzindo o incentivo da inovação tecnológica, cumpre ressaltar que o foco deveria ser no grau de risco trazido com a produção da decisão em questão, e no caso de um potencial alto e moderado, independente do tipo de empresa envolvida deveria ser obrigatória a revisão humana.

Poderia ser previsto algum tipo de incentivo tributário neste sentido em se tratando de startups, pois como explica Frank Pasquale<sup>2</sup>, o *accountability* impõe julgamento humano, “pois somente os homens podem cumprir a função crítica de ter certeza de que, na medida em que nossas relações sociais tornam-se cada vez mais automatizadas, aspectos de dominação e discriminação não são construídos invisivelmente dentro do código”. E continua, *verbis*:<sup>3</sup>

*“tais processos, ainda mais se forem totalmente automatizados e sem nenhum tipo de controle humano, podem ser fontes inesgotáveis de julgamentos equivocados e que reproduzam e intensifiquem ainda mais desigualdades e discriminações”.*

---

<sup>2</sup> The black box society: The secret algorithms that control money and information” p. 213.

<sup>3</sup> Ibidem, p. 108.

Ainda no caso de decisão automatizada o GDPR - Regulamento Geral de Proteção de Dados da União Europeia ao contrário da LGPD, traz também o DIREITO DE OPOSIÇÃO, ou seja, a possibilidade do titular dos dados opor-se à decisão gerada artificialmente, nos termos da Consideranda 71 e do artigo 22,1, quando afetar seus interesses.

O artigo 22, 1 preceitua que:

*“O titular dos dados tem o direito de não ficar sujeito a nenhuma decisão tomada exclusivamente com base no tratamento automatizado, incluindo a definição de perfis, que produza efeitos na sua esfera jurídica ou que o afete significativamente de forma similar.*

No lugar de se exigir o requisito da transparência apenas o que em certos casos se esbarra no problema da opacidade e caixa preta dos algoritmos, se fala em contestabilidade, e em contestability by design, assegurando-se uma explicabilidade que seja suficiente para permitir o direito de contestar ou rever determinada decisão tomada por uma IA. E em casos de ser impossível esta revisão de forma também automatizada, e não humana, por questões de limitação técnica, deverá ser obrigatória a revisão humana (<https://www.marcoalmada.com/publication/almada-direito-2020/almada-direito-2020.pdf>), lembrando também que diversos documentos internacionais colocam o controle humano da tecnologia como um fator determinante, como as diretrizes da OCDE. Observa-se que muitas vezes no âmbito de tal problemática estará a exigência de uma ponderação entre direitos fundamentais (Brkan, 2019), em aplicação, portanto, do princípio constitucional (e transnacional) da proporcionalidade.

Segundo a DECLARAÇÃO DE TORONTO (2018), deverá haver uma garantia de que grupos potencialmente afetados e especialistas sejam incluídos como atores com poderes decisórios sobre o design, e em fases de teste e revisão; revisão por especialistas independentes; divulgação de limitações conhecidas do sistema - por exemplo, medidas de confiança, cenários de falha conhecidos e limitações de uso apropriadas.

Por outro lado, o Algorithmic Accountability Act (EUA) traz a consideração acerca do denominado “high-risk automated decision system” – um sistema de decisão automática será considerado como de alto risco tendo em vista a novidade da tecnologia utilizada, sua natureza, escopo, contexto e propósito da decisão automatizada, que traga um risco significativo a privacidade, segurança ou resulte em decisões injustas e preconceituosas. Outro critério, é considerar se a decisão abrange áreas como consumidor, trabalho, saúde, economia, preferencias pessoais, interesses, comportamentos, de modo a alterar situações jurídicas do consumidor; monitoramento sistemático de uma área acessível ao público.

Há a obrigatoriedade da elaboração de relatório de impacto à proteção de dados, e de um relatório de impacto mais genérico, nos casos onde não há tratamento de dados, mas quando há o emprego de IA para automatização de processos de tomadas de decisão. Será então verificado e medido o impacto dos algoritmos em relação à sua precisão, equidade, discriminação, privacidade e segurança, aplicando-se às empresas de tecnologia com faturamento anual superior a U\$ 50.000.000 e que disponham de controle de dados pessoais de mais de 1.000.000 de consumidores ou mais de 1.000.000 de dispositivos.

Diante de algumas críticas a possibilidade de uma efetividade na prática da transparência, explicação e da revisão, pois estamos lidando em muitos casos com uma opacidade intrínseca ao sistema (*deep learning*), além de esbarrar em problemáticas outras como a do segredo comercial, seria interessante se pensar em ações afirmativas algorítmicas no combate à discriminação, no sentido de exigência de revisores independentes idôneos (justiça do algoritmo) e de se exigir a programação dos algoritmos para impedir a discriminação além de uma avaliação prévia dos dados utilizados para treinar os algoritmos.

Fala-se em *privacy by design e default* trazendo também na área da IA a incorporação das exigências legais na arquitetura do programa ou na explicação por design (*explanation by design*) ou explicação por padrão (*by default*) (Isabela Ferrari e Daniel Becker), com vistas a tentar trazer uma explicação sobre quais os critérios que foram utilizados e obrigação de documentação da origem dos bancos de dados e possíveis vieses. Contudo a proteção by design deveria voltar-se a todos os direitos fundamentais possivelmente afetados com tal tecnologia.

Não há também no PL qualquer previsão acerca de níveis de potencial dano das aplicações de IA ao contrário de diversas regulamentações da EU neste sentido, o que contribuiria para a abordagem baseada em risco, em risquificação e de coletivização, no sentido de uma tutela não apenas individual, mas coletiva, em um sistema de proteção proativo, abrangente e sistemicamente seguro, ou seja, uma efetiva proteção sistêmica, desde a concepção tecnológica, proporcionada pela *protection by design*, garantindo a *security by design*.

Outro ponto controvertido e crítico, que já desde a LGDP vem ocasionando interpretações equivocadas e que muitas vezes não se fundamentam no Constitucionalismo digital e no neoconstitucionalismo, é a temática da perícia de algoritmos de IA. Isto porque a própria LGPD dá margem a uma interpretação de que a proteção do segredo industrial, atrelado ao sigilo de algoritmos de inteligência artificial, seria um direito absoluto, ao invés de se pautar na necessidade inafastável da ponderação, tendo em vista a potencial colisão com outras normas de direitos fundamentais, assim compatibilizando-se com a necessidade de se postular pelos direitos à explicação, e revisão de decisões automatizadas.

No mesmo sentido, a preocupação de se interpretar e aplicar corretamente os princípios da razoabilidade e da proporcionalidade, pois muitas vezes são tidos como sinônimos, apesar de diversidade de origem, conceito e funções.

Segundo Wolfgang Hoffmann-Riem, a proteção judicial das pessoas adversamente afetadas, no caso de perícia em algoritmos de IA, poderia ser possibilitada pela introdução nos tribunais dos denominados procedimentos sigilosos; devendo as empresas revelar ao tribunal os algoritmos, em particular algoritmos que podem ser utilizados para pôr em perigo a liberdade - as máximas e os critérios em que se baseiam, a informação utilizada como input e, no caso dos sistemas de aprendizagem, as regras de formação utilizadas, se necessário também o tipo de utilização da análise de Big Data.

Contudo, essas informações não deverão tornar-se públicas e não deverão ser acessíveis às partes no processo, ou apenas o serão numa medida limitada, mas sim ao tribunal que aprecia os problemas, que pode, contudo, se necessário, mandar proceder a

um exame por peritos independentes (“Big Data e Inteligência Artificial: Desafios Para O Direito”).

Esclarece Wolfgang Hoffmann Riem que a transparência no design do algoritmo de IA não é apenas acerca da coleta e do processamento de dados, mas também do design tecnológico (o projeto técnico) e dos algoritmos utilizados em cada caso, dos sistemas algorítmicos. No entanto, a proteção dos segredos comerciais é contrária ao dever de divulgação.

É uma questão que demanda uma interpretação funcional e sistêmica, e não literal da legislação, voltando-se a uma adequada proteção dos direitos fundamentais envolvidos. Em seu entender, a quebra do segredo industrial seria justificada no caso de Direitos Fundamentais, em especial para evitar discriminação, estigmatização e seria justificada no caso de Direitos Fundamentais, em especial para evitar discriminação, estigmatização e manipulação. Ou no caso de existir outro interesse legítimo na divulgação equivalente à proteção de um segredo comercial. Se necessário, haveria proteção do segredo via procedimento sigiloso. Portanto, o segredo industrial não poderá ser interpretado e reconhecido como sendo um direito absoluto, mesmo porque a própria Lei de Propriedade Industrial abre exceções ao segredo industrial/de negócios, no caso de ações judiciais, desde que respeitado o segredo de justiça, devendo ser analisado o caso concreto mediante o procedimento de ponderação e aplicação da proporcionalidade (art. 206, da LPI). Apesar da previsão de ser possível uma auditoria pela ANPD no caso da LGPD (art. 20, p. 2º LGPD) para verificar os aspectos discriminatórios, quando se alegar o segredo industrial, verifica-se a fragilidade do dispositivo, a uma por se pautar apenas no aspecto da discriminação e não de

violação de outros direitos fundamentais, e também pela fragilidade institucional da ANPD por não ter total independente do Poder Executivo, segundo algumas críticas da doutrina, em comparação com o modelo europeu.

Diante da insegurança do processo judicial sob segredo de justiça no Brasil, já que não são raros os casos de publicação sem o respeito a tal limitação e de acesso por terceiros aos autos do processo, se maiores dificuldades, há que se pensar em uma proposta diante da realidade cultural brasileira, bem diversa do contexto europeu neste sentido, sendo uma alternativa viável a inversão do ônus da prova, proposta esta que parece estar de acordo com o procedimento da ponderação e aplicação da proporcionalidade, ou então reforçar-se a segurança no sentido de se garantir o segredo de justiça, seja como previsão de algum tipo de sanção ao funcionário que agir com negligência ao disponibilizar tal informação seja por via digital ou por via física, sem qualquer cuidado em termos do que razoavelmente se possa esperar.

Explica-se: ao invés de quebrar o segredo industrial, já que no Brasil não podemos garantir o segredo de justiça como suficiente para a preservação apenas entre as partes de tal revelação, poderia haver uma presunção de culpa caso a empresa se recusasse a informar tais dados relativos ao segredo industrial/de negócio, invertendo-se o ônus da causa, quando do outro lado da balança estiver presente um direito fundamental. Assim, ambos os Direitos Fundamentais estariam sendo protegidos mais adequadamente, sem o sacrifício total de qualquer deles (proporcionalidade em sentido estrito).

Uma interessante proposta é apontada por Wolfgang Hoffmann-Riem no sentido de se compatibilizar, a proteção aos



direitos fundamentais, princípios, responsabilização, e de outro lado, não impedir a inovação, denominada de “responsabilidade pela inovação”, ou “innovation forcing” (Hoffmann-Riem, Wolfgang. Teoria Geral do Direito Digital (pp. 13-14; p. 150 e ss.). Forense. Edição do Kindle.): trata-se da definição normativa de objetivos ou padrões que ainda não podem ser cumpridos sob o padrão de desenvolvimento atual, mas que são plausíveis de serem cumpridos. Tal direito concede então um período de implementação. Se ela expirar sem sucesso e não for estendida, o desenvolvimento e uso do tipo de IA em questão deve ser abandonado.

Já o Código de conduta dos engenheiros de robótica traz como medida para se garantir uma maior transparência a criação de caixas pretas preservando um log/ registro intangível de dados em relação às operações realizadas, envolvendo a lógica da produção da decisão automatizada. Seria a possibilidade de se inserir medidas técnicas no design dos sistemas algoritmos, visando, sobretudo, contribuir para sua compreensão, e explicação da tomada de decisões (outputs). Por fim, há também que ser destacada a importância do design participativo, desde que seja possibilitado o acesso aos meios necessários para eventuais prejudicados contestarem uma decisão, sendo que tal intervenção humana poderia se dar desde a etapa do design e desenvolvimento do sistema de IA.