

INTELIGÊNCIA ARTIFICIAL RESPONSÁVEL: SIGNIFICADOS E DESAFIOS¹

Wilson Engelmann²
DOI: 10.29327/5312711.1-1

Resumo: Os avanços observados com a Inteligência Artificial (IA) são constantes e crescentes. Se objetiva estudar os desafios que a IA está trazendo, notadamente para o campo regulatório. O problema a ser respondido é: como estruturar modelos de regulação para assegurar a Inteligência Artificial responsável? A pesquisa foi realizada a partir da pesquisa bibliográfica, no Portal de Periódicos da Capes, e documental, com preferência a publicações realizadas a partir de 2018, usando-se as palavras-chave abaixo indicadas. Como resultados se destaca a importância dos princípios jurídicos para a estruturação de modelos regulatórios que possam ter aplicação nacional e global, destacadamente a partir da comparação entre China e União Europeia. Também se verifica a importância da teoria

- 1 Este trabalho é o resultado parcial das pesquisas realizadas pelo autor no âmbito dos seguintes projetos de pesquisa: a) Chamada CNPq n. 09/2020 – Bolsas de Produtividade em Pesquisa – PQ, projeto intitulado: “Percurso para ressignificar a Teoria Geral das Fontes do Direito: o Sandbox regulatório como um elemento estruturante da comunicação reticular entre o Direito e as nanotecnologias”; b) “Sistema do Direito, novas tecnologias, globalização e o constitucionalismo contemporâneo: desafios e perspectivas”, Edital Fapergs/Capes 06/2018 – Programa de Internacionalização da Pós-Graduação no RS. Este trabalho também está vinculado às pesquisas realizadas pelo autor nos seguintes Centros Internacionais de Investigação: CEDIS – Centro de Investigação & Desenvolvimento sobre Direito e Sociedade, da Faculdade de Direito da Universidade Nova de Lisboa, Portugal; do Instituto Jurídico Portucalense, da Universidade Portucalense, cidade do Porto, Portugal; e do CEAD – Centro Francisco Suárez – Centro de Estudos Avançados em Direito da Universidade Lusófona de Lisboa, Portugal.
- 2 Doutor e Mestre em Direito Público, Programa de Pós-Graduação em Direito da Universidade do Vale do Rio dos Sinos – Unisinos, Brasil; realizou Estágio de Pós-Doutorado em Direito Público-Direitos Humanos, no Centro de Estudos de Seguridad (CESEG) da Universidade de Santiago de Compostela, Espanha; Professor e Pesquisador do Programa de Pós-Graduação em Direito – Mestrado e Doutorado e do Mestrado Profissional em Direito da Empresa e dos Negócios, ambos da Unisinos; Bolsista de Produtividade em Pesquisa do CNPq; Líder do Grupo de Pesquisa JusNano. ORCID: <https://orcid.org/0000-0002-0012-3559>. Lattes: <http://lattes.cnpq.br/7143561813892945> E-mail: wengelmann@unisinos.br.

sobre os Direitos Humanos, a partir de John Finnis, como referencial estruturante à arquitetura autorregulatória regulada e principiológica para os avanços e desafios trazidos pela IA, a fim de se promover as bases da IA responsável.

Palavras-chave: Inteligência Artificial. Autorregulação Regulada. Princípios. Direitos Naturais. Direitos Humanos. Direitos Fundamentais. Comparação principiológica entre China e União Europeia.

Sumário: 1. Introdução. 2. Conhecendo a Inteligência Artificial. 3. Um plano de vida coerente e a *phrónesis* aristotélica: estruturas metodológico-éticas para regular a IA. 4. Os princípios jurídicos como estruturas normativas para desenvolver a chamada IA responsável. 5. Considerações Finais. Referências.

1. Introdução

O desenvolvimento da Inteligência Artificial é um dado da realidade atual, com projeções bastante expressivas de crescimento e utilização nas variadas áreas de atuação do ser humano. Essa característica levanta uma preocupação: o ser humano será substituído em suas atividades pelos sistemas de Inteligência Artificial? Como se deverá regular esses avanços, assegurando-se o desenvolvimento da chamada Inteligência Artificial responsável? Esse o problema que o artigo pretende responder. O tema é atual e relevante dado o crescimento das variadas utilizações da IA, por um lado; e a ausência de uma regulação específica, por outro lado. Segundo estudo, elaborado pelo McKinsey Technology Trends Outlook (2022), publicado em agosto de 2022, existem quatorze tendências em IA que estarão em alta a partir de 2023, se destacando, por exemplo, a IA aplicada; conectividade avançada; futuro da bioengenharia; futuro da mobilidade e o futuro da energia limpa, dentre outras aplicações. Esse cenário promissor e desafiador, concomitantemente, sublinha a importância do tema e da discussão sobre a chamada IA responsável.

O objetivo geral do trabalho é: estudar os mecanismos para regular os avanços da IA responsável. Os objetivos específicos são: a) conhecer as características e os elementos estruturantes da IA; b) estudar, a partir de John Finnis, o denominado *plano de vida coerente* e a *phrónesis* aristotélica, como estruturas metodológico-éticas para regular a IA; c) conhecer os princípios jurídicos, publicados em documentos normativos pela China e União Europeia, como estruturas normativas para desenvolver a chamada IA responsável.

Metodologicamente o trabalho foi desenvolvido a partir da pesquisa bibliográfica, principalmente por meio de pesquisa realizada no Portal de Periódicos da Capes, acessado a partir do site principal da Biblioteca da Unisinos, e pesquisa documental. Todas as fontes foram selecionadas a partir da utilização das palavras-chave que apresentam este trabalho, priorizando-se publicações realizadas a partir de 2018.

Para se desenvolver os objetivos – geral e específicos – acima apresentados, o trabalho está dividido em três capítulos, conforme se poderá ler a seguir.

2. Conhecendo a Inteligência Artificial

A criatividade do ser humano poderá levá-lo à fronteira do conhecimento, gerando novidades e, simultaneamente, riscos imprevisíveis, talvez, no limite, incontroláveis. Essas perspectivas não significam que se deverá abondar o seu espírito inquieto e criativo. Pelo contrário, será preciso avançar com cuidado e respeitando certos balizadores que são apresentados pela própria natureza ao ser humano, como manter a saúde do corpo e mente, e preservar o meio ambiente em que se vive.

Uma dessas áreas se chama de “Inteligência Artificial” (IA, doravante), que se posiciona em um espaço de convergência da Computação Aplicada e outros quadrantes tecnológicos com domínios científicos do Direito e da Filosofia, principalmente, mas não exclusivamente. Se observa gradativamente que as áreas “duras” ou “exatas”, apesar de serem aquelas onde a IA se desenvolve e se torna possível, precisa da “ajuda” dos conhecimentos gerados nos domínios menos técnicos e

mais humanos ou vinculados às Ciências Sociais Aplicadas e Ciências Humanas.

Um exemplo desse cruzamento de áreas poderá ser colhido na época histórica próxima da pandemia da COVID 19. Vale dizer, um artigo publicado na *Revista Nature*, de 23 de março de 2021, dizia em seu título que a “Ciência não seria suficiente para salvar a humanidade”. Trazendo o conteúdo do artigo para a abordagem sobre a IA, se deve problematizar uma questão-chave: quem está sendo chamado para desenvolver e avaliar os riscos que poderão ser gerados pelo uso em larga escala da IA? Via de regra, se buscam conselhos e/ou pareceres de especialistas, ou seja, de pessoas em Ciência, Tecnologia, Engenharia e Matemática (a sigla dessas áreas em inglês é STEM) – apesar de estar claro, desde o início da projeção sobre IA, que o comportamento humano, as motivações e a cultura serão fundamentais para uma resposta eficaz. Essa abordagem precisa mudar. A Ciência viabiliza os sistemas de IA, mas as atividades ou áreas inseridas nas Ciências Sociais Aplicadas, Humanidades e Artes para as pessoas e a Economia (a sigla dessas áreas em inglês é SHAPE) ajudam a chegar às realidades sociais, além da análise e avaliação dos impactos sociais, éticos, jurídicos e econômicos dos avanços a partir da IA. A percepção da humanidade é mais robusta quando STEM e SHAPE se unem. Por isso, se poderá dizer que a IA, como um gênero, representa fenômenos sociais, biológicos e tecnológicos, concomitantemente. A partir desse cenário, não se deverá focar excessivamente em evidências de ensaios de controle randomizados, com as variáveis do funcionamento da IA controladas em laboratórios, mas levar em conta as evidências observacionais e qualitativas nas quais as Ciências SHAPE estão imersas (adaptado a partir de Shah, 2021).

Qual o fundamento desse cruzamento de conhecimentos das áreas STEM e SHAPE? Pelas diversas facetas que se encontra no próprio conceito de IA, que pode ser entendida como a

simulación de procesos de inteligencia humana por medio de máquinas que se extiende a través de capacidades tales como el reconocimiento de voz, la toma de decisiones, la búsqueda semántica y las diversas técnicas de aprendizaje automático (Devang et al, 2019).

Por meio do sistema de IA se poderá simular características próprias do ser humano, como raciocinar, conectar dados e projetar resultados e tomar decisões. Por isso, a IA também se define como a “[...] *capacidad de un sistema para interpretar correctamente los datos externos y aprender de ellos para utilizarlos con el fin de alcanzar objetivos específicos*” (Cuervo Sánchez, 2021). Todas essas características são muito próximas daquelas que os seres humanos também apresentam e que, até agora, distinguia-os dos demais seres vivos que habitam o Planeta Terra. E das máquinas também. Será que essa distinção estaria com os seus dias contados? Não se sabe, ainda.

E tem mais um detalhe: para que o sistema de IA possa desenvolver todas essas “competências” precisa do chamado “algoritmo”, que é uma coisa pequena e simples; uma regra usada para automatizar o tratamento de um dado (Barbosa, 2021). Uma espécie de receita: “se isso acontecer, então faça *b*; se não, então faça *c*”. Tal a lógica “se/então/mais” da computação clássica. Entretanto, essa é a concepção original, pois

[...] surgiu um significado mais ambíguo, indicando qualquer sistema de *software* de tomada de decisão complexo e grande; qualquer meio de obter um conjunto de dados de entrada – e avaliá-los rapidamente, de acordo com um determinado conjunto de critérios (ou “regras”) (Smith, 2018).

O sistema complexo que representa o algoritmo viabiliza a Inteligência Artificial (IA) por meio do armazenamento de dados e, a partir deles, consegue estruturar séries de dados, que no Direito podem gerar a “jurimetria”, por exemplo, como uma ferramenta para se ter uma ideia do modo como decide determinado tribunal, a partir de dados coletados por um longo período, que são tratados, estruturados e sistematizados, permitindo-se detectar tendências no sentido de julgamentos e temas escolhidos (Smith, 2018). A partir daí se pode referir que existe certa superioridade da máquina – do sistema – em relação ao ser humano? Ainda não se tem uma resposta acabada para essa pergunta, mas se observa uma tendência para uma resposta positiva.

Algoritmos não são novos. Por décadas, eles têm servido como componentes integrais de todos os programas de

computador. Na atualidade, os algoritmos, por meio dos avanços do aprendizado de máquina, estão criando uma sociedade amplamente automatizada, transformando muitas facetas da vida. Muitos produtos e serviços – incluindo filtros de e-mail como o “spam”, diagnósticos médicos, marketing de produto, e carros autônomos – dependem de algoritmos de aprendizado de máquina e sua capacidade de fornecer poder e velocidade surpreendentes. Os algoritmos atuais são “robôs” digitais que possuem habilidades efetivamente autônomas de adaptação e aprendizado (adaptado a partir de Smith, 2018).

3. *Um plano de vida coerente e a phrónesis aristotélica: estruturas metodológico-éticas para regular a IA*

Esse panorama abre diversas frentes para a pesquisa, especialmente para se chegar a um ponto de equilíbrio na denominada “IA responsável”. Como se poderá construir essa possibilidade? A seguir se levantam algumas questões que deverão ser levadas em consideração. Os computadores podem ser treinados para detectar padrões nos dados, mesmo padrões tão sutis que os humanos podem não perceber. Entretanto, quando se trata de identificar a relação entre causa e efeito, as máquinas geralmente ficam perdidas, pois elas carecem de uma compreensão de bom senso sobre o funcionamento do mundo, observável pelo ser humano na sua vida cotidiana. Para que os computadores realizem qualquer tipo de tomada de decisão, eles precisarão entender a causalidade. A incorporação de modelos de causa e efeito em algoritmos de aprendizado de máquina também pode ajudar máquinas autônomas móveis a tomar decisões sobre como navegar no mundo. Mesmo algo como o ChatGPT, o popular gerador de linguagem natural que produz texto que parece ter sido escrito por um humano, poderia se beneficiar da incorporação da causalidade. Com ela, o ChatGPT se poderia construir um plano coerente para o que estava tentando dizer e garantir que fosse consistente com os fatos, tais quais são conhecidos (adaptado a partir de Savage, 2023). A mescla da IA nesse novo chat também provocam desafios para as atividades jurídicas, pois trazem as chamadas “ferramentas de IA

generativas”. Essa nova categoria da IA está tendo um momento cultural e comercial, sendo apontada como o futuro da pesquisa, gerando disputas legais sobre direitos autorais e causando pânico em escolas e universidades. A tecnologia, que usa grandes conjuntos de dados para aprender a gerar imagens ou textos que pareçam naturais, pode ser uma boa opção para o setor jurídico, que depende fortemente de documentos e precedentes padronizados, como o caso dos contratos, além da geração automatizada de documentos jurídicos, que tem sido uma área de crescimento há décadas, mesmo em dias de tecnologia baseada em regras, pois os escritórios de advocacia podem recorrer a grande quantidade de modelos altamente padronizados e bancos de precedentes para estruturar a geração de documentos, tornando os resultados muito mais previsíveis do que com a maioria saídas de texto livre (Stokel-Walker, 2023).

Essas possibilidades parecem surpreendentes, pois a atividade da advocacia, como a elaboração de peças processuais e outros documentos, e a produção de textos, até agora, eram atividades intelectuais produzidas por seres humanos. A mobilização da linguagem e da criatividade, produziam contratos, peças processuais, textos para jornais, artigos científicos, dentre tantas outras produções intelecto-textuais. Se aprendia desde cedo na Escola a construção de diversas estágios para se aprimorar a escrita, a leitura e a compreensão textual (Floridi, 2023). Os avanços multifacetados das aplicações da IA geram curiosidade sobre quais novidades ainda virão. Paralelamente, se deverá investigar como regular tantas novidades e gerir os riscos que poderão decorrer desses avanços. Se observa uma euforia com as interações viabilizadas por meio da IA. Esse fenômeno poderá provocar um velamento dos impactos éticos, jurídicos e sociais.

Se observa que a IA sempre se vincula a um *processo de aprendizagem*, estruturada a partir de dados: quanto maior o número de dados que o sistema tem à sua disposição, maior o seu nível de aprendizado. Nesse cenário, cabe destacar a Proposta de Regulamento do Parlamento Europeu e do Conselho Europeu (chamada de *AI Act*) que apresenta a categoria conceitual de *sistema de inteligência artificial*, como um *software*, desen-

volvido a partir de um variado *conjunto de técnicas e abordagens* orientadas para um determinado conjunto de *objetivos definidos pelo ser humano*, gerando saídas como conteúdo, previsões, recomendações ou decisões que influenciam os ambientes com os quais eles interagem (Regulamento, 2021). Essa concepção evidencia a importância da participação do ser humano na definição dos objetivos que devem ser alcançados pelo sistema de IA. Se tem uma característica fundamental: *a tecnologia deverá ficar sob o controle humano* (Codagnone et al, 2022). Essa linha de preocupação é sublinhada por Luciano Floridi, quando refere: “A AI não é meramente outra utilidade que precisa ser regulada quando estiver madura. É uma força poderosa que está remodelando nossas vidas, interações e nossos ambientes” (Floridi, 2022). Pela análise do cenário atual do desenvolvimento da IA, se poderá dizer que ela está apenas em sua “infância”, ou seja, ainda existe espaço para avanços e não se poderá duvidar que essa referida remodelação de Floridi poderá se concretizar.

A garantia do “controle do ser humano” assegura a proteção dos direitos e garantias fundamentais (no plano regulatório interno), além de ser um elemento protetor para salvaguardar os Direitos Humanos (no plano regulatório externo). Se entende o seguinte percurso da tradição: os Direitos Naturais representaram o estágio inicial que preparou a emergência dos Direitos Humanos, que geraram, na sequência da caminhada histórica da humanidade, os Direitos Fundamentais, a partir da inserção dessa amálgama da tradição nos textos constitucionais. Aprofundando e qualificando ainda mais essa trajetória, se poderá dizer que os Direitos de Personalidade representam a versão contemporânea desses movimentos históricos de lutas e sofrimentos, assegurando a sua previsão no principal texto legal do Direito Privado, o Código Civil (Engelmann, 2023).

Para se desenhar um projeto regulatório adequado para o “mundo da IA”, caberá trazer algumas categorias filosóficas: uma delas, é o *exercício da prudência*, da medida e da cuidadosa avaliação das diversas facetas envolvidas na formulação das respostas jurídicas. Elas deverão ser perspectivadas com o manejo das características que compõe o *phrónimos*:

[...] pensa-se que é característico de uma pessoa que tenha *phrónesis* ser capaz de deliberar bem acerca do que é bom e conveniente para si mesma, não em relação a um aspecto particular [...] e sim acerca das espécies de coisas que nos levam a viver bem de um modo geral (Aristóteles, 2001, 1140a).

Para que o *phrónimos* possa alcançar a realização adequada, ou a melhor possível, dos bens humanos básicos (conhecimento, vida, sociabilidade, dentre outros definidos por John Finnis (2011) exigirá o chamado *plano de vida racional* que busca atingir e atender às características genéricas dos desejos e necessidades humanas e seu ciclo de ocorrência.

Esses aspectos, relativos a um plano de vida, inspiraram John Finnis na formulação de uma das exigências básicas da razoabilidade prática, que foi denominada de *plano de vida coerente*. Cada ser humano, e por relação direta – cada sociedade – deve ter um conjunto de propósitos e orientações, organizados com determinado grau de harmonia, passíveis de serem concretizados. Não se trata de planos imaginários e sem possibilidade de realização, mas de compromissos efetivos (Finnis, 2011, p. 103). A vida do ser humano deve estar norteadada por um projeto, alicerçado nas vivências pessoais passadas e presentes, mas lançando as vistas para o futuro. Vale dizer, é irracional que o ser humano apenas pense no imediato, esquecendo-se de amarrar cada momento, como um contínuo caminhar de cada um no contexto social.

O referido plano de vida, apesar de sua pensada coerência, poderá ser objeto de reformas ou revisões, onde se abandonam velhos projetos e se elegem novos. Os bens humanos estão integrados e interagem com o plano de vida desenvolvido por cada ser humano, justificando, portanto, alterações no decorrer de sua execução ou a partir do momento em que um novo plano é traçado, especialmente para atender a algum bem humano básico que necessite, naquele momento, uma maior atenção.

O *plano de vida racional* implica responsabilidade do ser humano, que deverá considerar-se tanto no momento de estipular as metas, como na projeção dos resultados das ações desencadeadas. Valendo também para a regulação da IA. O espaço público e o espaço privado se conectam a partir da introdução do conceito de *phrónesis*, que representa a tomada de

decisões não voltadas somente ao privado. Nesse particular, emerge um aspecto característico: como as pessoas não são iguais, os planos eleitos por cada uma também sofrem variações. Isso destaca a pessoalidade das escolhas, vinculadas a uma série de contornos individuais, marcados pela chamada experiência de vida, e que pode ser chamada de pré-compreensão, profundamente influenciada pela tradição onde o sujeito está imerso. A razoabilidade introduzida nesse contexto indica que o *plano de vida coerente* não está ligado apenas a uma concepção individual. Pelo contrário, deverá ser traçado com vistas ao coletivo, do qual o individual faz parte, mediante a projeção da concepção aristotélica de *phrónimos*.

O *plano de vida coerente*, desenhado a partir da tradição, e em observância aos contornos apresentados pelo mundo da vida, sugere um modelo de ação, para orientar a tomada de decisões sobre os possíveis riscos trazidos pela IA. Por conta disso, a formulação de um *plano de vida coerente* para o grupo social deverá ser agregada a uma outra *exigência básica da razoabilidade prática*, que John Finnis denomina de *bem comum*, justamente para reforçar a necessidade de planos coerentes para a vida do ser humano na sociedade, onde se conjugue o individual com o coletivo, o público com o privado, sem a primazia, inicial, de nenhum deles. O bem comum

se refere ao fator ou conjunto de fatores que, presentes como considerações no raciocínio prático de uma pessoa, atribui *um sentido a*, ou *uma razão para*, sua colaboração com os outros, desde o seu ponto de vista, estruturando uma razão para sua colaboração entre eles mesmos e com essa pessoa” (Finnis, 2011, p. 154).

Esse conjunto de elementos poderá auxiliar na construção do caminho rumo à IA responsável.

4. Os princípios jurídicos como estruturas normativas para desenvolver a chamada IA responsável

Uma das alternativas que se pode apresentar para a regulação da IA, orientada pela atitude do *phrónimos* e organizada

a partir de um *plano coerente de vida*, poderá se dar por meio de modelos regulatórios estruturados a partir de princípios, que se encontram vinculados diretamente com a referida amálgama da tradição que inicia com os Direitos Naturais, se conectam aos Direitos Humanos e desembocam nas Constituições contemporâneas, a partir do capítulo dos Direitos Fundamentais. Como se desenvolverá essa perspectiva? Na revisão de literatura sobre o tema se encontraram diversas fontes, dentre as quais se destacam Weng e Izumo (2019), Zhang et al (2022) e *AI Ethics Guidelines Global Inventory* (2020). Nessa última fonte, foram sistematizados mais de 160 documentos, estruturados a partir de princípios, e que foram elaborados por atores públicos e atores privados, localizados em países do Sul, do Norte, do Leste e de Oeste global. Essas fontes destacam o papel regulatório dos princípios.

Os princípios integram a categoria geral de “normas jurídicas”, ao lado das regras. Portanto, os princípios também apresentam uma carga de “dever ser”, com potencial diferente daquele carregado pelas regras. O trabalho jurídico, na resolução de problemas, com o uso dos princípios, que são considerados “mandados de otimização”, emprega critérios hermenêuticos distintos daqueles que norteiam a aplicação das regras, que são “mandados definitivos”. A perspectiva de “atribuição de sentido hermenêutico” exige um rigor metodológico, pois é orientada pelas consequências que a movimentação principiológica busca atingir (Engelmann, 2001; Engelmann, 2023). Se vislumbra um espaço crescente dos princípios e uma perspectiva hermenêutica renovada na sua aplicação, especialmente na construção de modelos normativos para regular os avanços da IA. Os princípios apresentam uma linguagem aberta para se amoldar a novas perspectivas sociais e que exigem regulação ágil e flexível. Tais não são as características da regulação legislativo-estatal. Por isso, se considera o alargamento do campo de aplicação dos princípios: não apenas os princípios expressa e implicitamente formulados nas Constituições ou em outros documentos normativos, mas também os princípios que orientam o intérprete, seguindo as linhas axiológicas fundamentais do ordena-

mento jurídico, na formulação da regra e na justificativa da solução a ser dada ao caso (Zaccaria, 2022). Se pode destacar que nesse trabalho de orientação do intérprete se conjugam os princípios inseridos no texto constitucional, assim como os princípios relativos aos Direitos naturais-humanos, abrindo a produção normativa para a conjugação de elementos internos com os elementos externos ao ordenamento jurídico de cada Estado.

Os princípios permitem introduzir no mundo jurídico uma série de regras que vêm da economia e das ciências, mas também da moral e outros campos do saber, sublinhando uma capacidade “normogenética ilimitada”; e, como tal, desempenham uma função essencial na difícil governança do pluralismo das fontes do Direito e da complexidade sociojurídica trazida pela IA devendo ser redefinida e recalibrada argumentativamente caso a caso pelo intérprete (Zaccaria, 2022, pp. 20-26).

O que sem dúvida se perde em termos de certeza e determinação do Direito, nessa revalorização dos princípios, ganha-se em troca de maior pluralismo: não mais apenas os legisladores, mas também (e sobretudo) os juízes, não mais apenas os Estados, mas também os novos sujeitos particulares e novas agências reguladoras; não mais apenas a legislação estadual, mas também a legislação supraestatal, internacional, comunitária, global, contribuem para configurar o conjunto do Direito (Zaccaria, 2022, pp. 26-29). Esse o cenário criativo, poroso e flexível, onde os princípios jurídicos atuam na atualidade, promovendo a criação normativa de estruturas típicas de *autorregulação*, em alguns momentos; em outros, de *autorregulação regulada* (Engelmann, 2022).

Esse é o percurso do processo atual de *mutação do Direito*, já não centrada mais preponderantemente na lei, mas nos dois polos da Constituição e dos documentos internacionais dos Direitos Humanos, por um lado; e dos tribunais nacionais e internacionais de Direitos Humanos, por outro. Isso gera, segundo Zaccaria (2022), um quadro de porosidade do Direito, que acentua dramaticamente a sua contingência e indeterminação e enfraquece a sua capacidade de ordenar e dar forma

à convivência civil. Com isso, surge naturalmente o delicado problema de distinguir o que é Direito, do que não é Direito, e para diferenciar novas formas de Direito de formas espúrias e degeneradas (Zaccaria, 2022, pp. 28-29).

Com esse panorama, se pode desenhar a revalorização normativa dos princípios, que conduzirão o referido processo de distinção, que tem a arquitetura projetada nas redes digitais, em um panorama específico e inédito da era da hiperconectividade (Floridi, 2015), que se caracteriza pela indiferença de se viver no modo on-line ou no modo off-line, pois as relações humanas e as demais dela decorrentes ocorrem no modo *OnLife*, ou seja, elas se desenvolvem espaço especial denominado *Infosfera*, onde tudo poderá ser perfeitamente analógico e digital, off-line e on-line (Floridi, 2018). Essa nova perspectiva que exige a mutação *do* e *no* Direito, inovando-se a produção e o reconhecimento do que é jurídico, daquilo que não é albergado pelo Direito, notadamente pelo reconhecimento de procedimentos regulatórios que levam a produção dos efeitos jurídicos para além do legalismo que sempre assume a forma de uma proposta de modelos rígidos, que garantem os bens da certeza e da previsibilidade, mas intrinsecamente inadequados para os tempos instáveis em que se vive (Grossi, 2020, pp. 79-85), gerados pela convergência tecnológica, dentre as quais, uma delas é a IA.

A referida revalorização normativa dos princípios se liga ao percurso histórico iniciado pelos Direitos Naturais, passando pelos Direitos Humanos até se chegar aos Direitos Fundamentais. Essa “caminhada” evidencia o papel da historicidade que se deverá projetar no “acontecer histórico” orientado no movimento do círculo hermenêutico (Gadamer). Essa perspectiva se encontra alicerçada na perspectiva trazida por Paolo Grossi (2020, pp. 86-87): “a historicidade significa plasticidade, vontade de se conformar por um devir constante, sem ter a pretensão de interrompê-lo, fixando-o em uma espécie de modelo atemporal e, portanto, anti-histórico”. Valorizar a historicidade é sublinhar o papel da tradição no *design* da atribuição de sentido hermenêutico ao *acontecer da linguagem* que compõe o ser humano. Vale dizer,

[...] Historicidade significa basicamente humildade diante da complexidade redescoberta, singeleza diante da mudança e do novo que ela comporta; historicidade significa consciência de que, se o resultado de um Direito autenticamente humano deve ser alcançado, ele não pode ser reduzido a uma vontade potestativa, imposta de cima, mas deve ser remetido à busca dos muitos recursos que o fluxo histórico traz consigo: variando e confirmando, ou seja, ressignificando o percurso (Grossi, 2020, pp. 86-98).

O percurso que liga os Direitos Naturais aos Direitos Fundamentais se enriquece dessa historicidade e reafirma a sua importância contemporânea, ao sustentar o plano de vida coerente da sociedade, na busca da regulação principiológica, iluminada pelos elementos estruturantes da *phrónesis*.

Se o Direito é fruto da invenção, então o cânone da historicidade do Direito, como abordagem inquestionavelmente afável da dimensão jurídica, representa a bússola orientadora para o legislador e para a comunidade de juristas/intérpretes, bem como o resgate de um Direito que é sempre e apenas experiência, ou seja, uma dimensão da vida (Grossi, 2020, p. 98). Os princípios representam essa conjugação da experiência de vida da sociedade. Por isso, eles estão em condições de, ao se renovarem, servirem para se estruturar modelos de *autorregulação regulada* para as novidades e os desafios gerados a partir da IA.

Colhendo-se um exemplo organizado pelo Fórum Econômico Mundial, que compara os princípios destacados no documento produzido pelo *Chinese National New Generation Artificial Intelligence Governance Professional Committee* (CNNGAI-GPC) e o documento elaborado pelo *European High-level Expert Group on AI* (HLEGAI), se poderá ter uma ideia das distinções existentes, mas, paralelamente, as convergências principiológicas, que confirmam o papel que os princípios representam na arquitetura da governança regulatória criativa para a IA. Esses aspectos evidenciam a possibilidade de se fazer um diálogo regulatório global a partir de uma fonte do Direito que deverá ser revalorizada: os princípios.

Quadro 1: Comparativo de princípios entre China e União Europeia

Princípios Éticos propostos pelo Chinese National New Generation Artificial Intelligence Governance Professional Committee	Exigências chave formulados pelo European High-level Expert Group on AI
1. Harmonia e amizade	1. Bem-estar social e ambiental
2. Equidade e Justiça	2. Diversidade, não-discriminação e equidade
3. Tolerância e compartilhamento	3. Supervisão humana
4. Respeito à privacidade	4. Privacidade e governança de dados
5. Segurança e controlabilidade dos sistemas de IA	5. Robustez técnica e segurança dos sistemas de IA
6. Responsabilidades compartilhadas	6. Transparência
7. Colaboração aberta	7. Responsabilidade
8. Governança ágil	

Fonte: World Economic Forum (2021).

Em muitos aspectos, os princípios chineses parecem semelhantes aos da União Europeia, promovendo justiça, robustez, privacidade, segurança e transparência. As perspectivas metodológicas, no entanto, revelam diferenças culturais claras. De qualquer modo, o percurso regulatório por meio dos princípios evidencia sinais de aproximação e diálogo.

As diretrizes chinesas derivam de uma perspectiva focada na comunidade e orientada a objetivos. “Um alto senso de responsabilidade social e autodisciplina” é esperado dos indivíduos para participar harmoniosamente em uma comunidade que promove tolerância, responsabilidades compartilhadas e colaboração aberta. Essa ênfase é claramente informada pelo valor confucionista de “harmonia”, como um equilíbrio ideal a ser alcançado através do controle de paixões extremas – os conflitos devem ser evitados. Além de uma advertência severa contra “uso ilegal de dados pessoais”, há pouco espaço para regulação estatal. A regulação não é o objetivo desses princípios, que são concebidos para orientar os desenvolvedores da IA da “maneira certa” para a elevação coletiva

da sociedade (World Economic Forum, 2021). Aqui se observa o caráter histórico da sociedade chinesa, que influencia a sua concepção sobre o papel dos princípios, ratificando o seu caráter de “mandados de otimização”, com as consequências jurídico-sociais estruturadas em cada caso concreto.

Os princípios europeus, emergindo de uma abordagem mais focada em direitos e baseados em direitos, expressam uma ambição diferente, enraizada no Iluminismo e coloridos pela história da Europa. Seu objetivo principal é proteger os indivíduos contra danos bem identificados. Enquanto os princípios chineses enfatizam a promoção de boas práticas, a UE se concentra na prevenção de consequências malignas. Fica bem clara a concepção entre meios e fins que os princípios deverão atender. A orientação chinesa atrai uma direção para o desenvolvimento da IA, de modo que contribui para a melhoria da sociedade, o modelo europeu define as limitações aos seus usos, para que não aconteça às custas de certas pessoas (World Economic Forum, 2021). A regulação estatal acaba trazendo elementos nacionais que dificilmente se comunicarão com a regulação de outros Estados. Por isso, a estrutura regulatória formada a partir dos princípios poderá gerar redes de normatização que aproximam os Estados, em uma verdadeira concepção global da IA responsável, ou seja, preocupada em compatibilizar o desenvolvimento da IA, mas com a centralidade da proteção do ser humano.

5. Considerações finais

O tema da IA ganha importância gradativa, especialmente pelos desenvolvimentos de pesquisa nas áreas exatas, especialmente de Ciência da Computação e Computação Aplicada. Se tem uma demora na construção da regulação desses avanços. Por isso, o problema de pesquisa que orientou este capítulo de livro foi: “Como se deverá regular esses avanços, assegurando-se o desenvolvimento da chamada Inteligência Artificial responsável?”.

A regulação da IA deverá buscar metodologias que consigam equilibrar o desenvolvimento da pesquisa e dos pro-

cessos de inovação decorrentes, com a forma e o tempo dessa regulação. O processo legislativo tradicional não parece ser o caminho mais adequado, dada a demora das suas variadas etapas. Por isso, se apresentou a possibilidade de construção de modelos de “autorregulação regulada”, onde se mescla a iniciativa privada, como um “novo” ator para regular a IA, com a observância de elementos que se encontram em panorama externo a esses atores, como o caso dos Direitos Humanos. Esses foram aqui desenhados a partir das contribuições de John Finnis, notadamente pela orientação encontrada em um dos “requerimentos ou requisitos da razoabilidade prática”, que é a estrutura de *um plano de vida coerente*: que é considerado uma possibilidade de conjugação de esforços e pressupostos de interesse público com o interesse privado. Tal conjugação, orientada pela arquitetura aristotélica da *phrónesis*, conduzirá à escolha de princípios, globalmente aceitos, para se estruturar modelos de “autorregulação regulada”.

A aceitação normativa dos princípios vem encontrando respaldo em alguns autores da Teoria Geral do Direito, conforme se estudou no capítulo, e, em escala crescente, eles vêm surgindo em diversos documentos de órgãos públicos e privados, em um elevado número de países. Esse fenômeno mostra que os princípios poderão ser um denominador regulatório comum para a IA promovendo um diálogo global de autorregulação regulada.

Não se tem, ainda, uma confirmação definitiva dessa possibilidade de resposta ao problema antes proposto, mas se observa ser um caminho que vai ganhando a adesão crescente, justamente pela perspectiva porosa dos princípios. Eles se abrem para a transição regulatória, a partir da orientação proveniente da hermenêutica filosófica, de orientação de Martin Heidegger e Hans-Georg Gadamer. A caminhada foi iniciada. Existe um largo espaço para o avanço e aprofundamento dessa pesquisa, sem se esquecer, também, da possibilidade de se testar esses modelos em laboratórios reais, como é o caso do *Living Lab Regulatório*. Aqui se trouxe uma pequena construção de uma pesquisa maior que o autor está desenvolvendo.

Referências

- AI Ethics Guidelines Global Inventory* (2020) [Online]. Available at: <https://inventory.algorithmwatch.org/about> (Accessed: 5 March 2023).
- Aristóteles (2001) *Ética a Nicômaco*. Translated from Greek, introduction and notes by Mário da Gama Kury. 4th edn. Brasília: UnB, 1140a.
- Barbosa, M. M. (2021) *Inteligência artificial: entre a utopia e a distopia, alguns problemas jurídicos*. Coimbra: GestLegal.
- Codagnone, C., Liva, G. and Rodriguez de Las Heras Ballell, T. (2022) *Identification and assessment of existing and draft EU legislation in the digital field: Study for the special committee on Artificial Intelligence in a Digital Age (AIDA)* [Online], Policy Department for Economic, Scientific and Quality of Life Policies, European Parliament, Luxembourg. Available at: [https://www.europarl.europa.eu/RegData/etudes/STUD/2022/703345/IPOL_STU\(2022\)703345_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2022/703345/IPOL_STU(2022)703345_EN.pdf) (Accessed: 6 March 2023).
- Cuervo Sánchez, C. A. (2021) Efectos de la inteligencia artificial en las estrategias de marketing: revisión de literatura. *Revista Internacional de Investigación en Comunicación aDResearch ESIC* (Primer semestre), 24(24), pp. 26-41 [online]. Available at: <https://doi.org/10.7263/adresic-024-02> (Accessed: 27 February 2023).
- Devang, V., Chintan, S., Gunjan, T. and Krupa, R. (2019) ‘Applications of artificial intelligence in marketing’. *Annals of the University Dunarea de Jos of Galati: Fascicle: XVII, Medicine*, 25(1), pp. 28-36, Galați, Romania.
- Engelmann, W. (2023) *Direito natural, ética e hermenêutica*. 2nd edn. revised, updated and expanded. Rio de Janeiro: Lumen Juris.

- Engelmann, W. (2022) 'Percursos para inovar a teoria geral das fontes do Direito: modelos de autorregulação regulada para as nanotecnologias, *sandbox regulatório* e princípios. Constituição, Sistemas Sociais e Hermenêutica' in Teixeira, A. V., Streck, L. L. and Rocha, L. S. (eds.) *Anuário do Programa de Pós-Graduação em Direito da Unisinos*, No. 18 [E-Book]. Blumenau, SC: Dom Modesto.
- Engelmann, W. (2001) *Crítica ao positivismo jurídico: princípios, regras e o conceito de Direito*. Porto Alegre: Sérgio Antonio Fabris.
- Finnis, J. (2011) *Natural law & natural rights*. 2nd edn. Oxford: Oxford University Press.
- Floridi, L. (2023) 'AI as agency without intelligence: on ChatGPT, large language models and other generative models'. *Philosophy and Technology* [online]. Available at: https://www.academia.edu/97068509/AI_as_Agency_without_Intelligence_on_ChatGPT_large_language_models_and_other_generative_models (Accessed: 5 March 2023).
- Floridi, L. (2022) 'The European Legislation on AI: a brief analysis of its philosophical approach' in Mökander, J. and Ziosi M. (eds.). *The 2021 Yearbook of the digital ethics lab, Digital Ethics Lab Yearbook*. Switzerland: Springer Nature.
- Floridi, L. (ed.) (2015) *The Onlife Manifesto: being human in a hyperconnected era*. London: Springer Open.
- Floridi, L. (2018) 'Soft Ethics and the Governance of the Digital', *Philosophy & Technology*, 31, pp. 1-8.
- Grossi, P. (2020) *Oltre la legalità*. Bari, Italy: Laterza & Figli.
- McKinsey Technology Trends Outlook (2022) *Applied AI* [Online] Available at: <https://www.mckinsey.com/~media/mckinsey/business%20functions/mckinsey%20digital/our%20in%20sights/the%20top%20trends%20in%20tech%202022/McKinsey-Tech-Trends-Outlook-2022-Applied-AI.pdf> (Accessed: 13 March 2023).

Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de Inteligência Artificial (Regulamento Inteligência Artificial) e altera determinados atos Legislativos da União Europeia (2021), Bruxelas, 21.4 COM(2021) 206 final. 2021/0106 (COD) [online]. Available at: https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0004.02/DOC_1&format=PDF (Accessed: 6 March 2023).

Savage, N. (2023) ‘Why artificial intelligence needs to understand consequences’, *Nature*, Outlook, 24 February [online]. Available at: <https://doi.org/10.1038/d41586-023-00577-1> (Accessed: 2 March 2023).

Shah, H. (2021) ‘COVID-19 recovery: Science isn’t enough to save us’. *Nature*, 591, p. 503 [online]. Available at: <https://doi.org/10.1038/d41586-021-00731-7> (Accessed: 27 February 2023).

Smith, A. (2018) ‘Franken-algorithms: the deadly consequences of unpredictable code’, *The Guardian*, 30 August [online]. Available at: <https://www.theguardian.com/technology/2018/aug/29/coding-algorithms-frankenalgos-program-danger> (Accessed: 27 February 2023).

Stokel-Walker, C. (2023) ‘Generative AI is coming for the lawyers’, *WIRED*, Business, 21 February [online]. Available at: <https://www.wired.com/story/chatgpt-generative-ai-is-coming-for-the-lawyers/> (Accessed: 3 March 2023).

Zaccaria, G. (2022) *Postdiritto: nuove fonti, nuove categorie*. Bologna: Mulino.

Zhang, D., Maslej, N., Brynjolfsson, E., Etchemendy, J., Lyons, T., Manyika, J., Ngo, H., Niebles, J. C., Sellitto, M., Sakhaee, E., Shoham, Y., Clark, J. and Perrault, R. (2022) *The AI Index 2022 Annual Report*, AI Index Steering Committee, Stanford Institute for Human-Centered AI, Stanford University, March [online]. Available at: https://aiindex.stanford.edu/wp-content/uploads/2022/03/2022-AI-Index-Report_Master.pdf (Accessed: 5 March 2023).

Weng, Y.-H. and Izumo, T. (2019) 'Natural Law and its implications for AI Governance', *Delphi* (Winter), 3, 10 December [online]. Available at: https://works.bepress.com/weng_yueh_hsuan/118/ (Accessed: 6 March 2023).

World Economic Forum, 2021. *Can China and Europe find common ground on AI ethics?* 19 November [online]. Available at: https://www.weforum.org/agenda/2021/11/can-china-and-europe-find-common-ground-on-ai-ethics/?utm_source=sfmc&utm_medium=email&utm_campaign=2762835_Agenda_weekly-26November2021&utm_term=&emailType=Agenda%20Weekly (Accessed: 8 March 2023).