

AUDITORIA COM BASE NA ÉTICA PARA O DESENVOLVIMENTO DE UMA INTELIGÊNCIA ARTIFICIAL CONFIÁVEL: UM MODELO DE *SANDBOX* COMO FERRAMENTA REGULATÓRIA ALICERÇADA EM PRINCÍPIOS

Wilson Engelmann

wengelmann@unisinors.br

Doutor e Mestre em Direito Público pelo Programa de Pós-Graduação em Direito (Mestrado e Doutorado) da Universidade do Vale do Rio dos Sinos – UNISINOS/RS/Brasil

Tainá Rafaela Bigaton

tainabigaton@gmail.com

Mestre em Direito pelo Programa de Pós-graduação em Direito da Universidade Comunitária da Região de Chapecó – UNOCHAPECÓ/SC/Brasil

A Inteligência Artificial (IA) é umas das maiores criações do homem e de acordo com Stephen Hawking (2018) o sucesso na criação da IA pode ser o maior acontecimento da história humana, mas, infelizmente, também pode ser o último, a menos que se aprenda a prevenir os riscos. Há uma lacuna de conhecimento entre aqueles que desenvolvem e aqueles que regulamentam a IA (KAUFMAN, 2021), e tal afirmação se soma a tendência de que esta assimetria informacional aumentará proporcionalmente ao crescimento da complexidade dos modelos. De todo modo, as questões abordadas com relação a IA, seus riscos e benefícios, apontam para uma transcendência teórica e territorial, e clamam por um olhar de perspectiva humanista, remetendo ao fato de que o Direito passou a ser confrontado por uma série de incertezas e complexidades, que precisarão de respostas criativas, sugeridas por meio de ferramentas diferentes daquelas tradicionalmente fornecidas pelo positivismo jurídico, em especial aquele de viés legalista. (ENGELMANN, 2013)

Com a finalidade de gerenciar os riscos e as preocupações que permeiam a relação HUMANIDADE x IA, e levando em conta que a legislação errada ou mal estruturada pode aniquilar com o desenvolvimento tecnológico, e se tratando de sistemas de IA e prevenção de riscos, é sabido que a regulação jurídica tende a ser mais eficiente em modo repressivo do que em modo preventivo, de maneira que o único caminho para alcançar uma regulação eficiente (preventivamente) no que tange aos sistemas de IA, é a aliança entre postulados éticos e um aumento da capacidade de auditoria técnica preventiva. (VERONESE, SILVEIRA, LEMOS, 2019) Para isso, a presente pesquisa buscou preencher essa lacuna de conhecimento, analisando a viabilidade e eficácia de uma auditoria baseada em ética como um mecanismo de

governança organizacional, voltando-se para uma Regulação do Futuro, pautada em princípios, que permitirá que as organizações operacionalizem seus compromissos éticos e validem as afirmações feitas sobre seus sistemas automatizados de tomada de decisão. (MÖKANDER, J., MORLEY, J., TADDEO, M. ET AL, 2021) A convergência principiológica encontrada quando analisados documentos internacionais e diretrizes éticas sobre a IA, como as do Japão e da União Europeia, dos documentos emitidos por órgãos multilaterais como a OCDE, e dos referenciais nacionais como a Resolução 332/20 do CNJ e a Estratégia Brasileira de Inteligência Artificial, demonstra que as preocupações relativas à IA giram sob o mesmo eixo: controle humano + desenvolvimento tecnológico = governança da IA. Princípios como o da transparência e explicabilidade norteiam debates sobre a utilização e os limites do desenvolvimento de sistemas de IA e frente a isso, esta pesquisa traz como objetivo geral a análise da viabilidade e da eficácia de um sistema *sandbox* regulatório, que atuará como uma ferramenta deontológica de auditoria mista. Como objetivos específicos estão a análise das restrições associadas a estes sistemas de auditoria e também, de como se dará a juridicização dos princípios que nortearão estas ferramentas deontológicas, a fim de que um Direito resiliente possa oferecer respostas sofisticadas para a sociedade. O trabalho foi desenvolvido a partir de uma estratégia de investigação que parte da Matriz Pragmático-Sistêmica na linha da Teoria dos Sistemas Sociais de Niklas Luhmann e de um método sistêmico-construtivista (ROCHA, 2015), com cunho exploratório, e são utilizadas técnicas de revisão bibliográfica e pesquisa documental.

A pesquisa demonstra a necessidade de uma abertura epistemológica do Direito a fim de recepcionar a IA e seus respectivos riscos e impactos éticos, legais e sociais (ELSI), a partir de ferramentas transdisciplinares para implementar políticas, procedimentos e padrões ao desenvolvimento de sistemas de IA. Pensar a inovação tecnológica e os sistemas de IA, como uma razão para regulação não é ignorar os demais racionais. É escolher as políticas e os procedimentos que permitirão que esses sistemas se desenvolvam e deem frutos benéficos à humanidade. Também, diante do fato de que a legislação errada ou mal estruturada pode aniquilar o desenvolvimento tecnológico, deve se considerar abster-se de regular por lei, mesmo que por um lapso temporal, e, desta forma, oportunizar que novos sistemas se desenvolvam dentro dos limites éticos estabelecidos para o estabelecimento de uma Governança da IA com o auxílio de ferramentas alternativas. Pensando nisso, esta pesquisa propõe o desenvolvimento de *AI Sandbox* regulatório, ou seja, uma ferramenta de auditoria

mista, pautada em princípios éticos e supervisão humana, na criação e utilização de ambientes seguros em que se submetam a testes os sistemas de IA, e se concentrou em compreender a viabilidade da construção de novas formas de regulação, que partam do controle social ao poder do Estado, e que por meio de auditorias e sistemas de segurança, possam garantir a conformidade com os princípios éticos e evitar um retrocesso no desenvolvimento tecnológico e todos os potenciais benefícios para a sociedade.

REFERÊNCIAS

ENGELMANN, Wilson; WERNER, Deivid Augusto. Inteligência artificial e Direito. In: FRAZÃO, Ana; MULHOLLAND, Caitlin. **Inteligência artificial e Direito: ética, regulação e responsabilidade**. 2. ed. rev., atual. e ampl. São Paulo: Revista dos Tribunais, 2020, p. 145-17

HAWKING, Stephen. **Breves respostas para grandes questões**. Trad. Cássio de Arantes Leite. 1 edição. Rio de Janeiro: Intrínseca, 2018.

Kaufman, D. (2021). Inteligência Artificial e os desafios éticos: a restrita aplicabilidade dos princípios gerais para nortear o ecossistema de IA. **PAULUS: Revista De Comunicação Da FAPCOM**, 5(9). <https://doi.org/10.31657/rcp.v5i9.453> (Original work published 30^o de junho de 2021)

Mökander, J., Morley, J., Taddeo, M. et al. Auditoria com base na ética de sistemas automatizados de tomada de decisão: natureza, escopo e limitações. **Sci Eng Ethics** 27, 44 (2021). <https://doi.org/10.1007/s11948-021-00319-4> Acesso em 15 jul 2021

ROCHA, Leonel Severo. Verbete “Niklas Luhmann”, in **Dicionário de Filosofia do Direito**. Coordenador: Vicente de Paulo Barreto. Rio de Janeiro: Renovar, 2006, p. 252

VERONESE, Alexandre; SILVEIRA, Alessandra; LEMOS; Amanda Nunes Lopes Espiñeira. Inteligência Artificial, mercado único digital e a postulação de um direito às inferências justas e razoáveis: uma questão jurídica entre a ética e a técnica. In: Ana Frazão; Caitlin Mulholland. (Org.). **Inteligência artificial e Direito: ética, regulação e responsabilidade**. 1ed.São Paulo: Thomson Reuters Brasil, 2019, v. 1, p.83-97