

Governança Ética Para Construção De Confiança Em Sistemas De Inteligência Artificial

Lucas Carini

Mestrando em Direito pela Faculdade Meridional – PPGD –IMED.
Email: lucas@guedesadvocacia.com.br

Fausto Santos de Moraes

Doutor em Direito Público – UNISINOS. Docente do PPGD da Faculdade Meridional - IMED/Passo Fundo.
Email: faustosmoraes@gmail.com

Resumo: Este artigo propõe a construção de um roteiro principiológico para servir como estrutura básica para guiar a governança ética aplicada a sistemas de inteligência artificial. Este estudo argumenta que um elemento-chave (necessário, mas não suficiente) na construção da confiança em Sistemas de Inteligência Artificial é a governança ética, ou seja, a existência de regulação. Dado o ritmo crescente da inovação, são necessários novos e ágeis processos de governança, para estabelecer diretrizes técnicas e legais. Nos últimos anos houve uma verdadeira proliferação de sistemas e produtos baseados em inteligência artificial, contudo, esta tecnologia está sendo utilizada, sem observância de princípios éticos o que dificulta a confiança pública nesses sistemas. Portanto, a conclusão, instrui para a construção de um roteiro principiológico, que orientará o desenvolvimento de sistemas de inteligência artificial para inovações responsáveis, seguindo uma regulamentação. O método utilizado foi o fenomenológico hermenêutico. A técnica de pesquisa foi a bibliográfica, em fontes secundárias.

Palavras-chave: Inteligência Artificial. Governança ética. Princípio. Regulamentação. Confiança.



UNIVERSIDADE FEDERAL DA PARAÍBA

João Pessoa, Programa de Pós-Graduação em Ciências Jurídicas

Governança Ética Para Construção De Confiança Em Sistemas De Inteligência Artificial

Lucas Carini

Fausto Santos de Morais

1 INTRODUÇÃO

O debate acerca da objetividade da pesquisa em ciências sociais – e, pois, sobre a própria natureza científica do conhecimento produzido nessa seara – está longe de perder a atualidade. Nos primórdios do racionalismo, a utilização de métodos de investigação próprios das ciências da natureza no âmbito das pesquisas elaboradas pelas ciências humanas figura como primeira tentativa de reabilitação do prestígio das ciências sociais por meio da busca do rigor metodológico nesse campo da pesquisa científica.

Em face da complexa matéria-prima das ciências humanas ou sociais, nestas a questão acerca dos limites e possibilidades do conhecimento científico formulada por Bacon e Descartes – principalmente no que pertine à objetividade – assoma com maior vigor. Nessa lógica, um dos aspectos problemáticos acerca da exigência de objetividade na pesquisa científica no âmbito das ciências sociais é a superação da ideologia enquanto obstáculo epistemológico.

Com efeito, a objetividade é o que garante o mais preciso conhecimento do objeto da pesquisa. Em ciências humanas ou ciências da vida, tal objeto é um recorte da vida humana nas áreas do

Direito, Sociologia, História, Economia etc. Nesse sentido, a identificação e neutralização da visão ideologicamente comprometida acerca da realidade é uma das mais importantes etapas para atingimento da medida de objetividade almejada pela ciência, visto que a ideologia elude o conhecimento do real independentemente da vontade do sujeito cognoscente.

Desde a negação do *status* científico às ciências sociais com fundamento na impossibilidade de produção de conhecimento objetivo - que caracteriza o desvio do racionalismo em cientificismo naturalista - até o enfoque epistemológico do tema empreendido por Karl Popper na *Lógica das Ciências Sociais*, o estatuto da objetividade científica tem evoluído com claras, mas pouco debatidas repercussões em matéria de construção de mecanismos para suplantar a ideologia enquanto obstáculo epistemológico.

Nessa ordem de ideias, problematiza-se neste trabalho a possibilidade de enfrentamento da ideologia na pesquisa científica em ciências sociais. Em razão da dificuldade maior que a questão alberga no âmbito das ciências humanas e da proeminência de Karl Popper na epistemologia contemporânea, faz-se um recorte da visão deste filósofo a respeito das ferramentas metodológicas aptas a combater a ideologia que compromete a objetividade no âmbito das ciências sociais.

Assim, o presente artigo utiliza o método dedutivo em pesquisa bibliográfica com o objetivo de analisar criticamente o pensamento de Karl Popper sobre objetividade e ideologia exposto na obra *Lógica das Ciências Sociais*, em face da questão sobre quais mecanismos podem ser utilizados na pesquisa científica elaborada no nicho das ciências humanas em ordem a superar o obstáculo epistemológico da ideologia e, destarte, atingir a medida de objetividade possível e necessária.

Para tanto, no tópico 1, introduz-se o tema da ideologia e as várias conceituações resultantes da evolução das investigações a tal respeito. No tópico 2, examina-se criticamente as teses descritas n'A

Lógica das Ciências Sociais de Popper. Por derradeiro, são enunciadas as considerações.

2 CONSEQUÊNCIAS IMEDIATAS E POTENCIAIS DO USO DE SISTEMAS DE INTELIGÊNCIA ARTIFICIAL SEM GOVERNANÇA ÉTICA

A possibilidade de um software desempenhar funções que supostamente eram exclusivamente humanas através de um sistema de Inteligência Artificial, tem trazido inúmeras aplicações nas mais diversas áreas. Utiliza-se como conceito de Inteligência Artificial a tecnologia que detecta automaticamente padrões nos dados e faz previsões com base nelas. É um método de análise inferencial que identifica correlações dentro de conjuntos de dados que podem, no caso de criação de perfil, ser usados como um indicador para classificar um assunto como um representante de uma categoria ou grupo (HILDEBRANDT 2008; SCHREURS *et al* 2008).

Ainda é possível fazer uma ampla distinção entre IA “estreita” e “geral” ou “ampla”. A IA estreita é uma aplicação de IA projetada lidar com uma tarefa específica, refletindo a maioria das aplicações de IA atualmente empregadas. Por outro lado, a IA geral ou ampla reflete a inteligência humana em sua versatilidade para lidar com tarefas diferentes ou gerais. (DALY *et al.* 2019, p. 5).

Outro conceito chave para o trabalho é a governança. Este conceito representa as respostas sobre quais os padrões éticos que devem orientar a IA. Essa questão encontra a sua relevância diante do caráter transnacional da tecnologia de IA, pois o desenvolvimento e a implementação da IA não se limita a determinados países ou jurisdições definidas territorialmente, mas funciona de forma globalizada em virtude da condição difusa que a Internet fez da comunicação. Nesse contexto, outros atores geopolíticos além dos

Estados acabam definindo a agenda ética e regulatória sobre a IA (DALY *et al.* 2019, p. 6).

Quanto ao termo governança especificamente, este recebe como definição geral o exercício da autoridade, controle, administração, poder de governo, ou seja, a maneira pela qual o poder é exercido. (WORLD BANK, 1992, p. 1-3). A acepção dada ao termo *governance* pelo dicionário Cambridge, como a maneira como organizações ou países são gerenciados ao mais alto nível (CAMBRIDGE DICTIONARY, 2019) completa a compreensão quanto ao seu sentido.

Tem-se utilizado a expressão *Digital Ethics* para tratar da relação entre a ética e a governança aplicados à AI, especialmente como o ramo que estuda e avalia os problemas morais relacionados aos dados, algoritmos e as correspondentes práticas. O objetivo principal é formular e apoiar soluções moralmente boas. Três têm sido os seus principais objetos de discussões: a ética dos dados, a ética de algoritmos e a ética das práticas. A ética dos dados analisa a geração, gravação, curadoria, processamento, disseminação, compartilhamento e uso de dados. Preocupa-se com os problemas morais colocados pela coleção, análise e aplicação de grandes conjuntos de dados. Os problemas variam desde o uso de grandes dados em pesquisa biomédica e ciências sociais para criação de perfis, publicidade e doação de dados e filantropia de dados, bem como dados abertos no governo projetos. (MAGRANI, 2019, p. 160).

Quanto a ética de algoritmos, deve imperar a responsabilidade algorítmica, de modo que, se faz necessário observar os algoritmos como objetos de criação e levar em consideração a intenção, inclusive de qualquer grupo, instituição, agência ou recursos humanos, ou seja, qualquer ator, que possam ter influenciado nos processos de design algorítmico. (MAGRANI, 2019, p. 160).

A ética das práticas, que esta relacionada com ética e governança, que por sua vez são aplicadas a AI, nos permite ver como os problemas relacionados a algoritmos pela sua natureza constitutiva, conhecimento e subjetividades, podem ser comandadas

por práticas de governança ética – ética das práticas. Tais práticas contêm alguma autoridade moral, pois impõem neutralidade e objetividade a um domínio que tenha relevância moral (a exemplo de um algoritmo projetado para identificar plágio). (MAGRANI, 2019, p. 159).

Transposta a breve definição quanto a terminologia, está claro que a Inteligência Artificial é uma tecnologia transformadora, com potencial para desafiar muitos aspectos regulatórios em curto, médio e longo prazo. A forma como o Direito e a poderá se adaptar aos avanços nesse campo depende de uma grande variedade de fatores sociais, culturais e econômicos, o que também influencia, por consequência, a forma como a jurisdição tratará sobre o assunto (STANFORD UNIVERSITY, 2016, p. 45).

De maneira geral o público tem confiança em Sistemas de Inteligência artificial, fato que os SIA são utilizados diariamente pelas pessoas, praticamente sem restrição. Seja no simples acesso à internet, através de buscadores que preveem as palavras antes de terminar de digitá-las, sistemas de *streaming* que sugerem músicas ou filmes de acordo com os gêneros anteriormente acessados pelo usuário, assistentes nos *smartphones* e também em versões para casa, que aprendem com as perguntas feitas diariamente ao ponto de que com apenas alguns dias de uso, esses assistentes são capazes de fazer sugestões antes do usuário perguntar, para não falar nos carros autônomos, cada vez mais comuns nas ruas de diversos países.

Enfim, os exemplos de produtos e serviços baseados em Sistemas de Inteligência Artificial, presentes no cotidiano das pessoas é extensa, o que demonstra que de maneira geral o público tem confiança nessa tecnologia, contudo, essa confiança é inspirada pelos benefícios que os usuários têm. É justamente nesse ponto que está o problema abordado no presente trabalho, a falta de regulamentação está deixando em aberto a lacuna das consequências que o uso de sistemas de Inteligência Artificial sem governança ética pode trazer.

Um exemplo singular das potenciais consequências do desenvolvimento e uso de SIA, sem governança ética, é o caso dos carros autônomos. Em um primeiro momento os benefícios de não precisar mais dirigir para se deslocar com seu veículo parece formidável, a possibilidade de utilizar esse tempo para outras atividades, como trabalhar, descansar ou até mesmo lazer, apresentam-se como benefícios insuperáveis.

Porém, ao analisar o que está incrustado no cerne dessa tecnologia, ou seja, seu código fonte, o seu algoritmo, é possível averiguar que existem consequências e potenciais consequências que o uso desse Sistema de Inteligência Artificial pode trazer.

No exemplo dos carros autônomos, podem ser inseridas no algoritmo do SIA, de acordo com o que se deseja obter, as diretrizes de segurança a serem empregadas. Que podem definir se o veículo, no caso de acidente, deva executar medidas que coloquem em risco os passageiros para priorizar salvar terceiros, como pedestres, ou se deve salvar os passageiros acima de qualquer outra hipótese.

Esse exemplo mostra o nível de complexidade com que a governança ética precisará lidar e evidencia a importância e a dificuldade da construção de regras para o desenvolvimento e uso de Sistemas e Inteligência Artificial.

Uma pesquisa realizada, justamente sobre a tomada de decisões dos carros autônomos, ilustra bem esse cenário. Em seis estudos a *Amazon Mechanical Turk*, descobriu-se que os participantes aprovaram veículos autônomos, no entanto, “prefeririam andar em veículos que protegessem os seus passageiros a qualquer custo, ao invés de carros autônomos que sacrificam seus passageiros para o bem maior”. (BONNEFON, RAHWAN, SHARIFF, 2016).

Mas, não é somente na área privada que os Sistemas de Inteligência Artificial, estão sendo empregados, as sociedades estão delegando cada vez mais processos complexos e de alto risco aos sistemas de IA. E essa realidade não está distante do Brasil. Há algum tempo o poder público brasileiro vem utilizando softwares com

funções estruturadas em IA. O Tribunal de Contas da União emprega três “robôs” para examinar editais de licitações, na busca de fraudes e irregularidades. Intitulados de Alice¹, Sofia e Monica, esses softwares, analisam e sugerem aprimoramentos em relatórios internos e acompanham todas as compras públicas. “Devido a esse experimento, outros órgãos de controle estão empregando essa tecnologia, como é o caso de Tribunais de Contas estaduais, Ministério Público e Polícia Federal e Controladoria Geral da União.” (FERRARI, BECKER, WOLKART, 2018, p. 3).

Outro órgão público a utilizar Sistemas de Inteligência Artificial é a Procuradoria-Geral do Distrito Federal, que utiliza um software privado para analisar o andamento a processos de execução fiscal (UNISINOS, 2017). A Receita Federal do Brasil, também anunciou o uso de sistemas de inteligência artificial no julgamento de processos, pelas suas delegacias, o software desempenharia em um primeiro momento o papel de ler os autos, identificar as alegações da defesa e, até mesmo, elaborar propostas de decisão. (OLIVON, PUPO, 2018).

Recentemente o Poder Judiciário, também aderiu a inteligência artificial, o Supremo Tribunal Federal anunciou a utilização do Victor, tido pela suprema corte, como o maior e mais complexo projeto de inteligência artificial do Poder Judiciário Brasileiro e talvez de toda a administração pública. O sistema está na fase de construção das redes neurais, que proporcionará aprendizagem a partir das decisões já proferidas pelo STF (2018).

Em outros países, o uso de Sistemas de Inteligência Artificial, já está causando consequências e discussões éticas. Como ocorreu no famoso caso *State v. Loomis, the Supreme Court of Wisconsin*, no qual “*Eric Loomis*, por furtar um veículo, evadir-se do agente de

¹ Alice é um acrônimo para Análise de Licitações e Editais, Sofia é um acrônimo para Sistema de Orientação sobre Fatos e Indícios para o Auditor, Sofia é um acrônimo para Sistema de Orientação sobre Fatos e Indícios para o Auditor, Sofia é um acrônimo para Sistema de Orientação sobre Fatos e Indícios para o Auditor. Informações foram extraídas do site da LegalLebs (2019), empresa desenvolvedora do Software Dra. Luiza.

trânsito e envolver-se em um tiroteio, foi condenado a seis anos de prisão, devido a antecedentes de agressão sexual, a pena aplicada não gerou surpresa (ISRANI, 2017)”.

O que chamou atenção foi o fato de que o juiz do caso utilizou como avaliação (rating) um software jurídico privado, intitulado de *COMPAS (Correctional Offender Management Profiling for Alternative Sanctions)*. Esse programa funcionava a partir de um algoritmo secreto, ao qual nem os juízes que o utilizam conheciam a sua programação. A sentença do caso foi proferida em 2013, *Loomis* recorreu à Suprema Corte Americana, que negou o pedido de admissão para julgamento, mesmo diante da apresentação do relatório elaborado pela *ONG Propublica*, que mostra o enviesamento do COMPAS contra afro-americanos. *Loomis* permanecerá preso até 2019 (FERRARI, BECKER, WOLKART, 2018, p. 02).

Os vários exemplos mencionados acima, mostram que a sociedade pode se beneficiar dessas novas tecnologias. Ao mesmo tempo, a IA pode ser mal utilizada ou comportar-se de maneiras imprevisíveis e potencialmente prejudiciais. Por isso a necessidade de construção de uma governança ética bem regulamentada para que a experiência tecnológica possa ser além de benéfica, confiável e segura.

3 INICIATIVAS DE GOVERNANÇA DE SISTEMAS DE INTELIGÊNCIA ARTIFICIAL

O debate que envolve todos os setores da sociedade – ou ao menos deveria envolver representantes da indústria, governos, acadêmicos e a sociedade civil – para pensar se as estruturas legais e regulatórias atuais são suficientes para governar os Sistemas de

Inteligência Artificial no seu desenvolvimento e uso, abrem caminho quanto a necessidade de abordagens técnicas e éticas a respeito

Todavia, essa questão não gira apenas na demanda regulatória, nem somente na discussão quanto a estruturas éticas. O impacto das tecnologias baseadas em SIA, envolvem questões mais abrangentes, que são o acúmulo de poder tecnológico, econômico e também político, nas mãos dos desenvolvedores de Sistemas de Inteligência Artificial.

Esse cenário, gerou uma verdadeira corrida pela liderança no segmento de inteligência artificial. Nos Estados Unidos, no ano de dois mil e dezenove, foi lançado o programa *American AI Initiative* (ORTEGA, 2019), com a intenção de concentrar incentivo à pesquisa para desenvolvimento da I.A.

O recente caso de incentivo à pesquisa em I.A. norte americano não é o único. Na América do Sul também estão surgindo iniciativas promissoras como a anunciada no último Fórum Econômico Mundial em Davos. Nesta oportunidade a Colômbia anunciou que fará grandes investimentos na área visando a inserção na chamada indústria 4.0. (ÉPOCA, 2019). O foco será o desenvolvimento de pesquisas em objetos como a Inteligência Artificial, Internet das Coisas e Blockchain.

Estima-se que atualmente mais de 18 países estão canalizando os seus esforços de pesquisa sobre inteligência artificial, se destacando países como França, México, Emirados Árabes Unidos e China. (DUTTON, 2019).

Mesmo diante desse cenário de intensificação das pesquisas para o desenvolvimento de Sistemas de Inteligência Artificial, ainda são escassas as iniciativas de governança e regulamentação. Contudo, é possível destacar alguns dos primeiros movimentos para estabelecer padrões de regulação técnica no desenvolvimento de SIA.

Um desses casos de padrões de regulamentação técnica no desenvolvimento de Sistemas de inteligência Artificial, é o caso dos padrões instituídos pela *International Organization for*

Standardization - ISO (2019)², como é o exemplo da ISO 13482, *Robots and robotic devices -- Safety requirements for personal care robots*, que estabelece requisitos de segurança para robôs de cuidados pessoais, que contém e são comandados por SIA embarcados (ISO, 2014).

O Subcomitê Técnico de Robôs e Dispositivos Robóticos da *British Standards Institution* (ISO, 2019a)³, publicou em abril de dois mil dezesseis, outro dos poucos exemplos de regulação para uma governança ética, de Sistemas de Inteligência Artificial. A norma BS 8611, tida como um guia para o projeto e aplicação ética de robôs e sistemas robóticos. “Esse regulamento, destaca-se por fornecer orientações éticas, para construção de um projeto seguro, visando a mitigação dos riscos sociais, como perda de privacidade, confiança, confidencialidade e o desemprego (BSI, 2016).”

Campolo endossa a ideia de regulação ou autorregulação ao sustentar que as empresas devem realizar testes rigorosos para garantir que os SIA não terão os seus vieses e erros amplificados em virtude de problemas com a utilização dos dados ou outras variáveis do programa. Para tal finalidade, o processamento dos dados deve ser documentado e disponibilizado publicamente, permitindo revisões e atualizações (CAMPOLO *et al.*, 2017).

Contudo, ainda são padrões de observância voluntária, não havendo nenhuma obrigatoriedade sobre adoção desses padrões de governança no desenvolvimento de Sistemas de Inteligência Artificial, como nos casos da robótica ou dos carros autônomos, por exemplo. Conforme visto no tópico anterior, a necessidade da construção de regulamentação está associada à preocupação de que

² A Organização Internacional para Padronização - ISO é uma organização internacional não governamental independente com uma adesão de 164 órgãos nacionais de normalização. Por meio de seus membros, reúne especialistas para compartilhar conhecimento e desenvolver Normas Internacionais relevantes, voluntárias, baseadas em consenso e relevantes para o mercado, que apoiem a inovação e forneçam soluções para os desafios globais (ISSO, 2019).

³ BSI - United Kingdom. A BSI é uma organização de distribuição sem fins lucrativos e oferece serviços globais nos campos vinculados de padronização, avaliação de sistemas, certificação de produtos, treinamento e serviços de consultoria (ISO, 2019a).

as tecnologias no ramo da Inteligência Artificial sejam desenvolvidas de forma a serem governáveis, abertas, transparentes e compreensíveis. Elas devem ser capazes de funcionar de forma eficiente com o público, mantendo o respeito aos valores estabelecidos em sociedade (UNITED STATES NATIONAL SCIENCE AND TECHNOLOGY COUNCIL, 2016, p. 39). Além disso, devem ser desenvolvidos planos de ação para que se incremente a diversidade dos ingressantes no mercado de trabalho nesse campo, incluindo pesquisadores e especialistas (UNITED STATES NATIONAL SCIENCE AND TECHNOLOGY COUNCIL, 2016, p. 28), a fim de construir uma estrutura de governança ética, com a finalidade de identificar os riscos e estar em constante aprimoramento das normas regulamentadoras dos SIA, a fim de mitigar os possíveis danos para a sociedade e potencializar os benefícios de sistemas de I.A.

4 O CAMINHO PARA A CONSTRUÇÃO DA GOVERNANÇA ÉTICA PARA SISTEMAS DE INTELIGÊNCIA ARTIFICIAL

O uso de Sistemas de Inteligência Artificial, seja de maneira direta, em softwares, como buscadores de internet, streaming de música, filmes e jogos, seja SIA na modalidade embarcada ou, como Robôs e carros autônomos, as aplicações da I.A. estão amplamente difundidas ao redor do mundo, o que sugere fortemente uma urgência voltada para a necessidade de regulação, através da construção de governança ética para os Sistemas de Inteligência Artificial.

Conforme demonstrado no decorrer desse trabalho, as decisões tomadas pelos Sistemas de Inteligência Artificial, têm consequências reais, para a segurança e bem-estar dos seres humanos; uma falha pode causar sérios danos, ferimentos ou colocar um inocente na prisão.

A difusão de produtos e serviços com SIA, mostra que as pessoas, de maneira geral tem confiança na Inteligência Artificial, pois, ela traz benefícios. A preocupação que fundamenta a construção de uma governança ética, é garantir a mitigação de consequências danosas às pessoas no uso dessa tecnologia. A construção de normas técnicas e legais, para instituir padrões de segurança e ética a serem seguidos no desenvolvimento de Sistemas de Inteligência Artificial, visa que a tecnologia além de benéfica seja segura.

É preciso construir a confiança pública através da regulação, o que coloca o governo estatal como um ator principal, responsável pela definição da agenda para o debate público. Colocadas as matérias relevantes para regulamentação, também será reservado ao governo estatal monitorar a aplicação das normas jurídicas produzidas. Para tanto, será necessária uma equipe qualificada e conhecedora do tema para a aplicação das normas e problematização de novas questões ainda não previstas.

A falta de regulamentação sobre a matéria aumenta a sensação de insegurança. Isto pode ser visto – inclusive por pesquisadores do calibre de Stephen Hawking – como um indício de que a IA pode ser uma real ameaça à humanidade (HAWKING, 2014).

A urgência na construção normas norteadoras ao desenvolvimento e ao uso de SIA aparecem por decorrência de outras pesquisas denunciando a impropriedade em programas de IA. Exemplo disso é o relatório do *Electronic Privacy Information Center*. Tal estudo apresentou que programas de IA semelhantes ao COMPAS estão sendo utilizados pela justiça criminal em vários estados norte americanos para a definição de fianças, predição de sentenças e definição do veredicto sobre o acusado. O problema está no fato de não haver testes de validação do algoritmo utilizado além das informações não estarem acessíveis ao público.

O papel da Governança ética é justamente atuar para que Sistemas de Inteligência Artificial, como o COMPAS, passem por um filtro normas integradas ao desenvolvimento dos produtos e serviços

com SIA, contribuindo para a construção da confiança na robótica e na inteligência artificial.

Contudo, nesse ponto, está um dos principais dilemas da construção de regulação da IA, quais princípios devem ser seguidos pelos desenvolvedores dessa tecnologia? Essa indagação expande a visão para a complexidade que está diante do desafio de regular os Sistemas de Inteligência Artificial. Isso porque não basta apenas entender como a tecnologia funciona, mas quais os protocolos devem ser adotados na projeção, desenvolvimento e operação das SIA.

Essa preocupação quanto aos valores e princípios que os produtos e serviços com SIA devem conter, revela a importância da construção de uma governança ética e, é explicitada quando relembramos o exemplo abordado no presente artigo, a respeito dos carros autônomos que salvam o passageiro a qualquer custo e os carros que sacrificam o passageiro para o bem maior, essas possibilidades de conduta é derivada da linguagem de programação, ou seja, da forma como o algoritmo do SIA embarcado nesses veículos foi construído. Condutas essas que revelam os princípios e valores éticos que foram incorporados na I.A. desses produtos.

Dentro do alcance e no limite temporal da pesquisa realizada para o presente artigo, não foi encontrada regulamentação no Brasil que englobasse de maneira completa princípios norteadores para o desenvolvimento de Sistemas de Inteligência Artificial, que compusesse a governança ética.

Os exemplos encontrados, são normas e orientações estrangeiras. Os já citados padrões técnicos, ISO 13482 (2014), que estabelece requisitos de segurança para robôs de cuidados pessoais, que contém e são comandados por SIA embarcados, instituído pela *International Organization for Standardization* – ISO e a norma BS 8611, publicada pela *British Standards Institution*, tida como um guia para o projeto e aplicação ética de robôs e sistemas robóticos.

Além desses poucos exemplos de regulação técnica para uma governança ética de Sistemas de Inteligência Artificial, é de singular importância destacar as iniciativas que estão surgindo na União

Europeia, que está demonstrando pioneirismo na construção de regulação de SIA.

O primeiro documento é a Resolução do Parlamento Europeu, de dezesseis de fevereiro de dois mil e dezessete, que contém recomendações à Comissão sobre disposições de Direito Civil sobre Robótica, documento que revela a preocupação com a regulação de serviços e produtos com SIA, conforme verifica-se no trecho abaixo colacionado:

Salienta que o desenvolvimento das tecnologias da robótica deve ser orientado para complementar as capacidades humanas, e não para as substituir; considera fundamental garantir que, no desenvolvimento da robótica e da inteligência artificial, os humanos tenham sempre o controlo sobre as máquinas inteligentes; considera que deve ser prestada particular atenção ao possível desenvolvimento de uma ligação emocional entre os seres humanos e os robôs, especialmente em grupos vulneráveis (crianças, idosos e pessoas com deficiência), e sublinha as questões suscitadas pelo grave impacto físico ou emocional que essa ligação emocional pode ter nos seres humanos; (PARLAMENTO EUROPEU, 2017).

Fica clara também a importância da construção da governança ética para regular os Sistemas de Inteligência Artificial na resolução aprovada pelo Parlamento Europeu, quando este discorre sobre a importância da construção de regulamentação frente as potenciais consequências que os produtos e os serviços com SIA podem trazer, frente a potencial ineficiência das normas atuais:

Observa que o reforço potencial de poderes através da utilização da robótica contrasta com um conjunto de tensões ou riscos e deve ser avaliado seriamente do ponto de vista da segurança, da saúde e da proteção humanas; da liberdade, da privacidade, da integridade e da dignidade; da autodeterminação e da não discriminação, e da proteção dos dados pessoais; Considera que o quadro jurídico em vigor da União deve ser atualizado e complementado, se for caso disso, através de princípios éticos que se coadunem com a complexidade da robótica e com as suas inúmeras implicações sociais, médicas e bioéticas; considera que um quadro ético orientador, claro, rigoroso e eficiente para a criação, a concepção, a produção, a utilização e a modificação de robôs é necessário para complementar as recomendações jurídicas do relatório e o acervo nacional

e da União existente; (PARLAMENTO EUROPEU, 2017).

O documento ainda direciona a atenção para o relevante papel que os princípios têm na composição da governança ética, que por sua vez, servirá de orientação basilar para o desenvolvimento de Sistemas de Inteligência Artificial, elencando alguns princípios que devem compor o quadro ético da regulação, sendo também seguidos no desenvolvimento de SIA:

Destaca que o quadro ético orientador deve basear-se nos princípios de beneficência, não-maleficência, autonomia e justiça, nos princípios e valores consagrados no artigo 2.º do Tratado da União Europeia e na Carta dos Direitos Fundamentais, tais como a dignidade do ser humano, a igualdade, a justiça e a equidade, a não discriminação, o consentimento esclarecido, o respeito pela vida privada e familiar e a proteção de dados, bem como em outros princípios e valores subjacentes do direito da União, como a não estigmatização, a transparência, a autonomia, a responsabilidade individual e a responsabilidade social, e em códigos e práticas éticas existentes; (PARLAMENTO EUROPEU, 2017).

No final de dois mil e dezoito, a União Europeia, em reunião plenária através da CEPEJ – Comissão Europeia para a Eficiência da Justiça adotou a Carta Ética para o Uso de Inteligência Artificial nos Sistemas Judiciais e seu Ambiente. (EUROPEAN COMMISSION FOR THE EFFICIENCY OF JUSTICE (CEPEJ), 2018, p. 07).

O documento da CEPEJ forja logo na sua introdução os cinco princípios éticos para o uso de Sistemas de Inteligência Artificial no poder judiciário, embora a aplicação das disposições da carta seja recomendada tanto para atores públicos como para atores privados.

Como primeiro princípio o respeito aos direitos fundamentais, no qual a concepção e implementação de ferramentas de inteligências artificial precisam ser compatíveis com os direitos fundamentais⁴. Em segundo lugar a carta enumera o princípio da não discriminação,

⁴ “Principle of respect for fundamental rights: ensure that the design and implementation of artificial intelligence tools and services are compatible with fundamental rights” (Cf. Ibid., Pág. 7).

que visa prevenir a criação ou reforço de discriminações entre indivíduos ou grupos. (EUROPEAN COMMISSION FOR THE EFFICIENCY OF JUSTICE (CEPEJ), 2018, p. 07).

Como terceiro princípio é elencado o a qualidade e segurança, aplicado ao tratamento das decisões judiciais e dados judiciais, para que sejam utilizadas fontes certificadas em um ambiente tecnológico seguro. (EUROPEAN COMMISSION FOR THE EFFICIENCY OF JUSTICE (CEPEJ), 2018, p. 07). Em quarto lugar está o princípio da transparência, neutralidade e integridade intelectual, que basicamente poderia ser dividido em três, todavia foi aglutina em apenas um, esse princípio busca tornar acessível e compreensível as metodologias de processamento de dados, inclusive autorizando auditorias externas. (EUROPEAN COMMISSION FOR THE EFFICIENCY OF JUSTICE (CEPEJ), 2018, p. 07). Por fim, está consignado o princípio do controle pelo usuário, princípio que permite ao usuário ser um ator esclarecido e não apenas um terceiro afastado sem informação.

Os princípios trazidos pelos documentos europeus, demonstram coerência com a construção da governança ética de Sistemas de Inteligência Artificial. Tanto a Carta Ética para o Uso de Inteligência Artificial nos Sistemas Judiciais e seu Ambiente, como a Resolução do Parlamento Europeu com recomendações à Comissão sobre disposições de Direito Civil sobre Robótica.

Outra iniciativa vanguardista, surgiu na cidade de Nova York, em 2018, onde foi promulgada a primeira lei de responsabilidade algorítmica dos Estados Unidos. A lei cria uma força-tarefa que tem a incumbência de investigar os “sistemas de decisão automatizados”, o termo “significa implementações computadorizadas de algoritmos, incluindo aquelas derivadas de aprendizado de máquina ou outras técnicas de processamento de dados ou inteligência artificial, usadas para tomar ou ajudar na tomada de decisões.”⁵ (NEW YORK CITY COUNCIL, Law Number 2018/048).

⁵ Tradução livre, texto original em inglês: “The term “automated decision system” means computerized implementations of algorithms, including those derived from machine learning or other data processing or artificial intelligence techniques, which are used to make or assist in making decisions.”

A lei é voltada para os sistemas utilizados pelo poder público, no caso, pelas agências governamentais, para tomada ou auxílio de decisões que impactam o público. A intenção da legislação é criar um grupo de especialistas para que identifiquem os sistemas automatizados de decisão e os seus impactos desproporcionais ou prejudiciais para a população, nos termos da lei, colaciona-se os procedimentos recomendados para criação:

(a) Critérios para identificar quais sistemas de decisão automatizados da agência devem estar sujeitos a um ou mais dos procedimentos recomendados por essa força-tarefa nos termos deste parágrafo; (b) Desenvolvimento e implementação de um procedimento através do qual uma pessoa afetada por uma decisão relativa a uma regra, política ou ação implementada pela cidade, onde tal decisão foi tomada por ou com a assistência de um sistema automatizado de decisão da agência, possa solicitar e receber uma explicação dessa decisão e a base da mesma; (c) Desenvolvimento e implementação de um procedimento que possa ser usado pela cidade para determinar se um sistema de decisão automatizado da agência afeta desproporcionalmente pessoas com base em idade, raça, credo, cor, religião, origem nacional, gênero, deficiência, estado civil, parceria status, status de cuidador, orientação sexual, alienação ou status de cidadania; (d) Desenvolvimento e implementação de um procedimento para tratar de casos em que uma pessoa é prejudicada por um sistema automatizado de decisão da agência, se for considerado que um sistema desse tipo afeta desproporcionalmente pessoas com base em uma categoria descrita na alínea (c); (e) Desenvolvimento e implementação de um processo para disponibilizar publicamente as informações que, para cada sistema de decisão automatizado de cada agência, permitirá ao público avaliar significativamente como esse sistema funciona e é usado pela cidade, inclusive disponibilizando publicamente informações técnicas sobre esse sistema, onde apropriado; e (f) A viabilidade do desenvolvimento e implementação de um procedimento para arquivar sistemas de decisão automatizados da agência, dados usados para determinar relações preditivas entre dados para esses sistemas e dados de entrada para tais sistemas, desde que isso não precise incluir sistemas de decisão automatizados da agência que cessaram sendo usado pela cidade antes da data efetiva desta lei local. ⁶(NEW YORK CITY COUNCIL, Law Number 2018/048).

⁶ Tradução livre, texto original em inglês: (a) Criteria for identifying which agency automated decision systems should be subject to one or more of the procedures recommended by such task force pursuant to this paragraph; (b) Development and implementation of a procedure through which a person affected by a decision concerning a rule, policy or action implemented by the city, where such decision was made by or with the assistance of an agency automated decision system, may request and receive an explanation of such decision and the basis therefor; (c)

A lei de Nova Iorque – Lei nº 2018/048 (*NEW YORK CITY COUNCIL, Law Number 2018/048*), tem como objetivos norteadores a “justiça, responsabilidade e transparência”, conforme denota-se no trecho colacionado acima. Tais valores conversam com os princípios elencados pelo documento da União Europeia.

Essas são importantes iniciativas para solucionar o problema da falta de regulação e conduzem para a construção de uma governança ética dos sistemas de inteligência artificial e reduzir o problema da insegurança jurídica gerada, sobretudo nos sistemas aplicados ao judiciário. Nesse sentido, um dos valores ou princípios estruturantes para esse combate, é o aumento da transparência. Nas palavras de Bernard Harcourt, professor de direito da Universidade de Columbia que estudou as avaliações de risco, considera que “a lei de Nova York é um primeiro passo importante, mas que a verdadeira solução para um problema cada vez maior é o aumento da transparência”⁷. (BERNARD, 2017).

Os documentos citados apresentam uma grande quantidade de pontos de convergência entre os princípios, apontando o caminho para construção da Governança éticas de Sistemas de

Development and implementation of a procedure that may be used by the city to determine whether an agency automated decision system disproportionately impacts persons based upon age, race, creed, color, religion, national origin, gender, disability, marital status, partnership status, caregiver status, sexual orientation, alienage or citizenship status; (d) Development and implementation of a procedure for addressing instances in which a person is harmed by an agency automated decision system if any such system is found to disproportionately impact persons based upon a category described in subparagraph (c); (e) Development and implementation of a process for making information publicly available that, for each agency automated decision system, will allow the public to meaningfully assess how such system functions and is used by the city, including making technical information about such system publicly available where appropriate; and, (f) The feasibility of the development and implementation of a procedure for archiving agency automated decision systems, data used to determine predictive relationships among data for such systems and input data for such systems, provided that this need not include agency automated decision systems that ceased being used by the city before the effective date of this local law. (NEW YORK CITY COUNCIL, Law Number 2018/048).

⁷ Tradução livre, texto original em inglês: “*algorithmic bias in government agencies is widespread and growing, especially in areas like policing and criminal adjudication that are getting cannibalized by the facile solution of predictive tools.*”

Inteligência Artificial. A fim de congregar regras e princípios, como por exemplo, não causar danos às pessoas, sejam danos físicos ou morais (como preconceitos ou discriminação), respeitar os direitos fundamentais (como por exemplo liberdade, privacidade), que incluem dignidade da pessoa humana, privacidade (proteção de dados), também garantir a transparência e acesso desses dados, para assegurar a responsabilidade. Ou seja, os Sistemas de Inteligência Artificial precisam seguir os princípios da governança ética, mitigando as consequências negativas, construindo confiança pública.

5 CONSIDERAÇÕES FINAIS

O desenvolvimento e uso de Sistemas de Inteligência Artificial ainda carece de muita pesquisa a fim de embasar as discussões políticas, jurídicas e legislativas sobre o tema. A aplicação dessa tecnologia na sociedade de maneira indiscriminada, sem ponderação e medida das suas implicações, pode acarretar graves violações aos Direitos Fundamentais e danos reais para as pessoas.

Assim, a construção da Governança ética de Sistemas de Inteligência Artificial, mostra-se necessária. Do mesmo modo, importante salientar que o presente trabalho não tem a pretensão de se opor a tecnologia da inteligência artificial, pelo contrário, o propósito é de contribuir para harmonizar os algoritmos com a interação humana.

Por outro lado, a construção da Governança ética de Sistemas de Inteligência Artificial tem inúmeros desafios para serem superados. Um dos principais é o problema da formalização e implementação de valores éticos e o raciocínio moral em Sistemas de Inteligência Artificial, problemática essa, que por sua abrangência, está fora dos limites do presente trabalho. Desse modo, o próximo

passo será fornecer aos SIA uma espécie de governador ético, que deverá reproduzir as normas éticas, legais e padrões técnicos predeterminados pela sociedade ou pelo país onde serão utilizados. Nesse sentido, o SIA poderá avaliar as consequências de suas ações de acordo com o conjunto de regras com as quais foi programado.

Embora existam muitas questões a serem desenvolvidas para que a Governança Ética de Sistemas de Inteligência Artificial torne-se realidade prática, não há dúvidas de que esta temática precisa ser aprimorada diante da necessidade de regulamentação de SAI em consideração às potenciais consequências jurídicas de uma desregulamentação.

Data de Submissão: 30/09/2019

Data de Aprovação: 03/12/2019

Processo de Avaliação: *double blind peer review*

Editor Geral: Jailton Macena de Araújo

Editor de Área: Jailton Macena de Araújo

Assistente Editorial: Matheus Victor Sousa Soares

REFERÊNCIAS

BERNARD, Zoë. The first bill to examine 'algorithmic bias' in government agencies has just passed in New York City. **Business Insider**. Dez 19, 2017 às 16:03. Disponível em: <https://www.businessinsider.com/algorithmic-bias-accountability-bill-passes-in-new-york-city-2017-12?IR=T>. Acesso em: 21 nov. 2019.

BONNEFON, Jean-François. RAHWAN, Iyad. SHARIFF, Azim. The social dilemma of autonomous vehicles. **Science**, 24 Jun. 2016. Vol. 352, Issue 6293, pp. 1573-1576. DOI: 10.1126/science.aaf2654.

Disponível em:

<https://science.sciencemag.org/content/352/6293/1573>. Acesso em: 22 jun. 2019.

BRASIL. **LEI Nº 13.709**, DE 14 DE AGOSTO DE 2018. Planalto. Disponível em http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Lei/L13709.htm. Acesso em: 08 jun.2019.

BRASIL. **Senado Federal**. Disponível em:

<http://www.senado.gov.br/noticias/Jornal/emdiscussao/motos/saude/estudo-da-organizacao-mundial-da-saude-oms-sobre-mortes-por-acidentes-de-transito-em-178-paises-e-base-para-decada-de-aco-es-para-seguranca.aspx>. Acesso em: 02 abr.2019.

BROEKAERT, Kris. ESPINEL, Vitória A. **How can policy keep pace with the Fourth Industrial Revolution?** 13 de fevereiro de 2018. Disponível em:

<https://www.weforum.org/agenda/2018/02/can-policy-keep-pace-with-fourth-industrial-revolution/>. Acesso em: 22 jun. 2019.

BSI. **Standard highlighting the ethical hazards of robots is published**. 14 de abril de 2016 .Disponível em:

<https://www.bsigroup.com/en-GB/about-bsi/media-centre/press-releases/2016/april/-Standard--highlighting-the-ethical-hazards-of-robots-is-published/>. Acesso em 23 jun. 2019.

BSI. United Kingdom. **International Organization for Standardization**. Disponível em:

<https://www.iso.org/member/2064.html>. Acesso em: 23 jun. 2019.

CAMPOLO et al. **AI Now 2017** Report. p. 1-2. Disponível em:

<https://assets.contentful.com/8wprhhvnpfco/1A9c3ZTCZa2KEYM64Wsc2a/8636557c5fb-14f2b74b2be64c3ce0c78/_AI_Now_Institute_2017_Report_.pdf>. Acesso em: 26 nov. 2019.

DALY, Angela; Hagendorff, Thilo; Hui, Li; Monique Mann, Vidushi Marda, Ben Wagner, Wei Wang and Saskia Witteborn. **Artificial Intelligence Governance and Ethics: Global Perspectives**. Report 28 June 2019 - The authors acknowledge the research assistance of Ms Jing Bei and Mr Sunny Ka Long Chan, and the comments and observations from participants in the CUHK Law Global Governance of AI and Ethics workshop, 20-21 jun. 2019.

DUTTON, Tim. **Construindo um mundo AI**: Relatório sobre estratégias nacionais e regionais de IA.. Disponível em:

<https://www.cifar.ca/cifarnews/2018/12/06/building-an-ai-world-report-on-national-and-regional-ai-strategies>. Acesso em: 18 jun. 2019.

EPIC.ORG. **Algorithms in the Criminal Justice System.**

Electronic Privacy Information Center. Disponível em:

[<https://epic.org/algorithmic-transparency/crim-justice/>]. Acesso em: 26 nov. 2019.

ÉPOCA. Negócios on-line. **Presidente da Colômbia anuncia novo polo de indústria 4.0 no país.** Disponível em:

<https://epocanegocios.globo.com/Mundo/noticia/2019/01/presidente-da-colombia-anuncia-novo-polo-de-industria-40.html>. Acesso em: 18 jun. 2019.

EUROPEAN COMMISSION FOR THE EFFICIENCY OF JUSTICE (CEPEJ) **European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment.** Pág. 7. Disponível em <https://www.coe.int/en/web/cepej/cepej-european-ethical-charter-on-the-use-of-artificial-intelligence-ai-in-judicial-systems-and-their-environment>. Acesso em: 08 jun. 2019.

FERRARI, Isabela; Becker, Daniel; Wolkart, Erik Navarro. **ARBITRIUM EX MACHINA: PANORAMA, RISCOS E A NECESSIDADE DE REGULAÇÃO DAS DECISÕES INFORMADAS POR ALGORITMOS.** Arbitrium ex machina: framework, risks and the governance of decisions informed by algorithms Revista dos Tribunais | vol. 995/2018 | Set / 2018 DTR\2018\18341.

GOVERNANCE. **In: Cambridge Dictionary.** Cambridge

University Press 2019. Disponível em:

<https://dictionary.cambridge.org/pt/dicionario/ingles/governance>. Acesso em: 23 nov. 2019.

HAWKING, Stephen; RUSSELL, Stuart; WILCZEK, Max Tegmark; Frank. **Transcendence looks at the implications of artificial intelligence - but are we taking AI seriously enough?** 2014.

Disponível em:

<<http://www.independent.co.uk/news/science/stephen-hawking-transcendence-looks-atthe-implications-of-artificial-intelligence-but-are-we-taking-9313474.html>>. Acesso em: 23 nov. 2019.

HILDEBRANDT, M. (2008). **Defining profiling:** A new type of knowledge? In M. Hildebrandt & S. Gutwirth (Eds.), Profiling the European citizen: Cross-disciplinary perspectives (pp.17-45). Dordrecht: Springer.

ISO. ISO 13482:2014. **Robots and robotic devices - Safety requirements for personal care robots.** International Organization For Standardization, 2014. Disponível em:

<https://www.iso.org/standard/53820.html>. Acesso em: 23 jun. 2019.

ISO. **About us.** International Organization For Standardization.

2019. Disponível em: <https://www.iso.org/about-us.html>. Acesso em 23.06.2019.

ISO. **Members:** membership – memberbody. International. Organization For Standardization. 2019a. Disponível em: <https://www.iso.org/member/2064.html>. Acesso em 23.06.2019.

ISRANI, Ellora. **Algorithmic due process:** mistaken accountability and attribution in State v. Loomis. JOLTdigest. 2017. Disponível em: <https://jolt.law.harvard.edu/digest/algorithmic-due-process-mistaken-accountability-and-attribution-in-state-v-loomis-1>. Acesso em: 03 jun. 2019.

LEGALLEBS, **empresa desenvolvedora do Software Dra. Luiza.** Disponível no site: <http://www.draluzia.com/>. Acesso em: 24 jun. 2019.

MAGRANI, Eduardo. **Governance Of Internet Of Things And Ethics Of Artificial Intelligence A Governança Da Internet Das Coisas E A Ética Da Inteligência Artificial** Revista Direitos Culturais | Santo Ângelo | v. 13 | n. 31 | p. 153-190 | set./dez. 2018. DOI: <http://dx.doi.org/10.20912/rdc.v13i31.2816>.

NEW YORK CITY COUNCIL. **Introduction of Law Number 2018/048.** Disponível em: <https://legistar.council.nyc.gov/legislationDetail.aspx?ID=3137815&GUID=437A6A6D-62E1-47E2-9C42-461253F9C6D0>. Acesso em: 21 nov. 2019.

OLIVON, Beatriz e PUPO, Fábio. **Receita testa inteligência artificial em julgamentos.** Revista Valor Economico. 2018. Disponível em: <https://www.valor.com.br/legislacao/5473055/receita-testa-inteligencia-artificial-em-julgamentos>. Acesso em: 01 jun. 2019.

ORTEGA, João. **Trump anuncia iniciativa para promover inteligência artificial nos EUA.** Disponível em: <https://www.startse.com/noticia/nova-economia/60827/trump-inteligencia-artificial-nos-eua>. Acesso em: 18 jun. 2019.

PARLAMENTO EUROPEU. **Disposições de Direito Civil sobre Robótica Resolução do Parlamento Europeu**, de 16 de fevereiro de 2017, que contém recomendações à Comissão sobre disposições de Direito Civil sobre Robótica (2015/2103(INL)). Disponível em: <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML+TA+P8-TA-2017-0051+0+DOC+PDF+Vo//PT>. Acesso em: 29 jun. 2019.

STANFORD UNIVERSITY. **Artificial Intelligence And Life In 2030: One Hundred Year Study On Artificial Intelligence (AI100) Report**, 2016. p. 45-49. Disponível em: https://ai100.stanford.edu/sites/default/files/ai100report10032016fnl_singles.pdf. Acesso em: 29 nov. 2019.

SUPERIOR TRIBUNAL FEDERAL. **Inteligência artificial vai agilizar a tramitação de processos no STF.** 30 maio 2018.

Disponível em:

<http://www.stf.jus.br/portal/cms/verNoticiaDetalhe.asp?idConteudo=380038>. Acesso em 24 jun. 2019.

SUPERIOR TRIBUNAL FEDERAL. **Ministra Cármen Lúcia anuncia início de funcionamento do Projeto Victor, de inteligência artificial.** Disponível em:

<http://www.stf.jus.br/portal/cms/verNoticiaDetalhe.asp?idConteudo=388443>. Acesso 24 jun. 2019.

SUPERIOR TRIBUNAL FEDERAL. **Projeto VICTOR do STF é apresentado em congresso internacional sobre tecnologia.**

Disponível em:

<http://portal.stf.jus.br/noticias/verNoticiaDetalhe.asp?idConteudo=390818>. Acesso em 24 jun. 2019.

UNIÃO EUROPEIA. **GDPR - General Data Protection**

Regulation. Disponível em: <https://eugdpr.org/>. Acesso em: 08 jun. 2019.

UNISINOS, Instituto Humanitas. **Dr^a Luzia, primeira robô-advogada do Brasil, já tem trabalho pela frente.** 2017.

Disponível em: <http://www.ihu.unisinos.br/78-noticias/569427-dr-luzia-primeira-robo-advogada-do-brasil-ja-tem-trabalho-pela-frente>. Acesso em: 20 jun. 2019.

UNITED STATES NATIONAL SCIENCE AND TECHNOLOGY COUNCIL. **Preparing for the Future of Artificial Intelligence,** 2016. p. 2-40. Disponível em:

<https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf>. Acesso em: 26 nov. 2019.

Ethical Governance For Building Trust In Artificial Intelligence Systems

Lucas Carini

Fausto Santos de Morais

Abstract: In this article, it is proposed the construction of a theoretical guide to serve as a basic structure to guide ethical governance applied to artificial intelligence systems. This article argues that a key element (necessary, but not sufficient) in building trust in Artificial Intelligence systems is ethical governance ie the existence of regulation. Given the increasing pace of innovation, new and streamlined governance processes are needed to establish technical and legal guidelines. In recent years there has been a real proliferation of systems based on artificial intelligence, however, this technology is being used, without observing ethical principles which makes public confidence in these systems difficult. Therefore, the construction of a theoretical guide will guide the development of artificial intelligence systems as responsible innovations, following standards and regulations. The method used was deductive. The research technique was the bibliographical one, in secondary sources.

Keywords: Artificial Intelligence. Ethical Governance. Principle. Regulation. Trust.