

ISSN 2526-0774

Vol. IV
Jan - Dez 2020



Recebido: 14.10.2019
 Aceito: 12.12.2019
 Publicado: 30.01.2020

¹ Universidade Federal de Juiz de Fora.
Juiz de Fora, Minas Gerais – Brasil.
<https://orcid.org/0000-0001-9810-8164>

² Universidade Federal de Juiz de Fora.
Juiz de Fora, Minas Gerais – Brasil.
<https://orcid.org/0000-0003-4760-5149>

³ Universidade Federal de Juiz de Fora.
Juiz de Fora, Minas Gerais – Brasil.
<http://orcid.org/0000-0002-0495-3218>

⁴ Universidade Federal de Juiz de Fora.
Juiz de Fora, Minas Gerais – Brasil.

⁵ Universidade Federal de Juiz de Fora.
Juiz de Fora, Minas Gerais – Brasil.

⁶ Universidade Federal de Juiz de Fora.
Juiz de Fora, Minas Gerais – Brasil.

⁷ Universidade Federal de Juiz de Fora.
Juiz de Fora, Minas Gerais – Brasil.

INTELIGÊNCIA ARTIFICIAL E DIREITOS HUMANOS: IMPACTOS E DILEMAS ÉTICOS ATUAIS

ARTIFICIAL INTELLIGENCE AND HUMAN RIGHTS: CURRENT IMPACTS AND ETHICAL DILEMMAS

*Ana Carolina de Souza Pereira*¹

*Ana Luisa Brêtas Bruno*²

*Ariane Meirelles de Azevedo*³

*Caroline da Rosa Pinheiro*⁴

*Lívia Melo Honorato Campos*⁵

*Nicole Vieira Abreu Ornellas*⁶

*Victória Presoti Paixão*⁷

Resumo

O presente artigo visa analisar, por meio de uma revisão bibliográfica e estudo de caso exploratório, alguns dilemas éticos impostos pela Inteligência Artificial e suas consequências no âmbito dos direitos humanos. O problema de pesquisa possui como finalidade investigar os dilemas éticos surgidos a partir do uso de sistemas de Inteligência Artificial e como estes se relacionam com violações de direitos humanos, tendo como prisma as diretrizes éticas acordadas pela União Europeia no ano de 2019. A partir das diretrizes, desenvolveu-se uma pesquisa de cunho documental e bibliográfico com uma análise qualitativa. Entre os resultados encontra-se a necessidade de uma regulamentação dos sistemas de Inteligência Artificial, que seja capaz de normatizar questões, as quais possuem impacto nos direitos humanos, a fim de mitigar seus potenciais riscos de violação e o desafio incutido a essa tarefa, tendo em vista os possíveis impactos inibitórios de uma regulamentação à inovação tecnológica.

Palavras-chave

Direitos humanos. Diretrizes éticas. Inteligência Artificial. União Europeia. Inovação.

Abstract

The present article aims to analyze, through a bibliographic review, some ethical dilemmas imposed by Artificial Intelligence and its consequences in the sphere of human rights. The research problem aims to investigate the ethical dilemmas arising from the use of Artificial Intelligence systems and how these relate to human rights violations, taking as a perspective the Ethical Guidelines agreed by the European Union in the year 2019. Based on the guidelines, a documentary and bibliographic research and a qualitative analysis were developed. Among the results is the need for regulation of Artificial Intelligence systems, which is able to regulate issues, that have an impact on Human Rights, in order to mitigate their potential violation risks and the challenge instilled in this task in view of the possible inhibitory impacts of regulation on technological innovation.

Keywords

Human Rights. Ethical guidelines. Artificial Intelligence. European Union. Innovation.

1. INTRODUÇÃO

O presente artigo tem por objetivo refletir sobre algumas discussões éticas decorrentes do crescimento usual de sistemas de Inteligência Artificial em escala global e suas possíveis consequências para os direitos humanos¹. Por conta dos inúmeros conceitos acerca do termo ética, o presente artigo se baseou na definição proposta pelo Ethics Guidelines for Trustworthy AI, elaborado pela União Europeia, com a finalidade de regular as questões relativas a esse avanço tecnológico (ETHICS GUIDELINES FOR TRUSTWORTHY AI, 2019b, p.37):

A ética é uma disciplina acadêmica que constitui um subdomínio da filosofia. Em termos gerais, trata de questões como "O que é uma boa ação?", "Qual é o valor de uma vida humana?", "O que é a justiça?" ou "O que é uma boa vida?". Na ética acadêmica, há quatro grandes domínios de investigação: i) metaética, que diz sobretudo respeito ao significado e à referência das frases normativas, e à questão de saber como os seus valores de verdade podem ser determinados (caso existam), ii) ética normativa, um meio prático para determinar a orientação moral a seguir, mediante um exame das normas de boa e má ação e da atribuição de um valor a ações específicas, iii) ética descritiva, que visa fazer uma investigação empírica do comportamento e das convicções morais das pessoas, iv) ética aplicada, respeitante ao que somos obrigados (ou autorizados) a fazer numa situação específica (muitas vezes historicamente nova) ou num determinado domínio (muitas vezes sem precedentes históricos) de possibilidades de ação. A ética aplicada trata de situações da vida real, em que as decisões têm de ser tomadas sob pressão do tempo e muitas vezes com uma racionalidade limitada. A ética da IA é geralmente encarada como um exemplo de ética aplicada e centra-se nas questões normativas suscitadas pela concessão, pelo desenvolvimento, pela implantação e pela utilização da inteligência artificial. No âmbito dos debates éticos, os termos "moral" e "ético" são frequentemente utilizados. O termo "moral" refere-se aos padrões de comportamento concretos e factuais, aos costumes e convenções que podem ser encontrados em culturas, grupos ou indivíduos específicos num determinado momento. O termo "ético" refere-se a uma avaliação valorativa das ações concretas e dos comportamentos de uma perspectiva sistemática e acadêmica.

Sendo assim, o trabalho está dividido nos seguintes tópicos: o item 2 apresenta um breve histórico sobre a evolução da Inteligência Artificial. O item 3 discorre acerca do conceito formulado pela União Europeia sobre Inteligência Artificial e seu funcionamento. O item 4 aborda as Diretrizes Éticas elaboradas pelo grupo de especialistas da União Europeia. Os itens 5 e 6 abordam algumas violações de direitos humanos já ocorridas, devido ao uso desregulado dos sistemas de Inteligência Artificial. Por fim, o item 7 expõe as conclusões encontradas a partir do estudo realizado e aponta possíveis soluções sobre o conturbado tema.

O marco teórico adotado são as diretrizes da União Europeia, formuladas por uma comissão de especialistas no assunto, no ano de 2019. Em relação à técnica metodológica empregada, optou-se pela realização de uma revisão bibliográfica, sob o viés do estudo de caso, utilizando a ferramenta de generalização analítica por meio da forma de estudo de caso exploratório, de acordo com a metodologia de Yin.

A escolha deste marco teórico se justifica devido a sua capacidade de articulação dos elementos jurídicos à complexidade encontrada no tema tratado. A técnica metodológica empregada

¹O conceito de Direitos Humanos adotado é o mesmo disciplinado pela Declaração Universal dos Direitos Humanos, ou seja, de direitos universais, inerentes a todos os seres humanos, independentemente de quaisquer condições. A declaração pode ser acessada pelo seguinte endereço eletrônico: <https://nacoesunidas.org/direitoshumanos/declaracao/>

tem o objetivo de trazer especificidades dos casos concretos e sua relação com as diretrizes éticas formuladas.

O objetivo da pesquisa é analisar como os direitos humanos, em uma “Era de Inovações Tecnológicas”, podem ser respeitados, levando em consideração a difícil tarefa de construção de Inteligência Artificial confiável e ética. Cabe ainda apresentar o problema que guiou este estudo: como os direitos humanos são afetados a partir do uso cada vez mais crescente de sistemas de Inteligência Artificial, tendo em vista os dilemas éticos que circundam o tema, principalmente a criação de uma regulamentação que funcione como um mecanismo de solução em casos concretos de violações de direitos humanos e não um freio à inovação tecnológica.

Considerando o nível da discussão, a importância do tema se revela a partir de sua atualidade e da carência de regulamentação. Diante do exposto, resta necessário compreender os pormenores dessa inovação bem como o seu impacto nos Direitos Humanos.

2. INTELIGÊNCIA ARTIFICIAL: PORMENORIZANDO A COMPREENSÃO

O início da compreensão sobre o tema deve partir da noção de que a Inteligência Artificial usa principalmente a racionalidade, ou seja, refere-se à capacidade de decidir qual melhor ação tomar para se atingir um determinado objetivo, tendo em conta certos critérios de otimização e recursos disponíveis.

Outro ponto, o qual deve ser observado em relação a definição do que seria uma Inteligência Artificial, é a não possibilidade desta ser analisada de forma isolada, sendo tido como um sistema, uma vez que, frequentemente, a IA é incorporada como componente de sistemas maiores. Contudo, para fins práticos, será utilizado nesse artigo o termo “Inteligência Artificial”, sua abreviação (IA), bem como o termo “sistemas de Inteligência Artificial”. (HIGH-LEVEL EXPERT GROUP ON AI, 2019a, p. 1).

3. AFINAL, O QUE É INTELIGÊNCIA ARTIFICIAL?

A Convenção Europeia sobre IA, propôs uma definição do que seria Inteligência Artificial, delineando melhor o que é possível se entender conceitualmente a seu respeito:

A Inteligência Artificial (IA) refere-se a sistemas que apresentam um comportamento inteligente, analisando o seu ambiente e tomando medidas – com algum grau de autonomia – para atingir objetivos específicos.

No entanto, esta concepção foi já foi expandida pela comissão de alto nível da União Europeia (HIGH-LEVEL EXPERT GROUP ON AI, 2019b, p. 1), objetivando abarcar certos aspectos que possam alcançar um entendimento comum a noção de IA. Sendo assim, é proposto o uso de uma definição em que:

Sistemas de IA são sistemas de software (e possivelmente hardware) projetados por seres humanos que, dado um objetivo complexo, atuam na dimensão física ou digital, percebendo seu ambiente através da aquisição de dados estruturados ou não estruturados coletados, raciocinando sobre o conhecimento, ou processando as informações, derivadas desses dados e decidindo a melhor ação a ser tomada para atingir o objetivo dado. Os sistemas de IA podem usar regras simbólicas ou aprender um modelo numérico, e também podem adaptar seu comportamento analisando como o ambiente é afetado por suas ações anteriores.

3.1 MACHINE LEARNING: COMO O SISTEMA DE INTELIGÊNCIA ARTIFICIAL OPERA

Como disciplina científica, a IA inclui diversas abordagens e técnicas, sendo que hoje a maioria desses sistemas detém mecanismos de raciocínio (que inclui planejamento, programação, representação, busca e otimização), robótica (que insere características de controle, percepção, sensores, bem como a integração de todas as outras técnicas em sistemas cyberfísicos) e autoaprendizagem, intitulado de Machine learning. Este último configura uma forma de incorporar algoritmos de aprendizado em máquinas, permitindo que o computador transforme essas informações em experiências novas dentro dos seus limites estruturais e as incorporem ao seu já imenso conjunto de dados, fazendo com que esses sejam processados e sua reação ao comando seja aprimorada a cada interação. (HIGH-LEVEL EXPERT GROUP ON AI, 2019a, p. 4)

Existem várias formas de autoaprendizagem, entre elas está a chamada aprendizagem profunda (deep learning), a qual se caracteriza pela detenção por parte da IA de uma rede neural composta por múltiplos estratos que possibilitam aprender uma relação global de input-output em etapas sucessivas, tornando o sistema mais autônomo e com menos necessidade da orientação humana. Outro tipo de abordagem é intitulado de aprendizado de reforço, no qual o sistema IA possui certa liberdade para tomar suas decisões e, a partir delas, se institui um método de recompensa de acordo com uma boa ou má decisão. (HIGH-LEVEL EXPERT GROUP ON AI, 2019a, p. 4)

A partir dos elementos discutidos acerca do machine learning, e seus tipos derivados, é estabelecida uma discussão em volta dos limites éticos que devem ser postos para que os sistemas de Inteligência Artificial não prejudiquem o desenvolvimento da humanidade.

O dilema toma uma forma mais delineada quando, a partir da inserção de diversos dados em uma máquina programada, com algoritmo de aprendizado em setores distintos e sem qualquer discernimento ético previamente programado, esta é capaz de realizar um cruzamento de dados, calcular índices estatísticos e fornecer uma reação que pode afetar consideravelmente os direitos humanos, inerentes aos indivíduos. Por exemplo, se os dados de treinamento forem tendenciosos, isto é, se não forem equilibrados ou inclusivos o suficiente, o sistema de IA formatado e tais dados não serão capazes de generalizar de modo eficaz e, possivelmente, tomarão decisões injustas que podem favorecer alguns grupos em detrimento de outros.

Afinal, se os dados produzidos pela IA são reflexos de comportamentos humanos, e muitas das atitudes humanas revelam-se preconceituosas e atentatórias a direitos, há grande probabilidade de reprodução pelas máquinas destas posturas. Desse modo, torna-se claro que o acompanhamento humano sempre será imprescindível, mesmo em ambientes altamente autônomos, para que exista o discernimento humano na forma de controle de qualidade dos resultados produzidos e reproduzidos por máquinas.

3.2 A RELEVÂNCIA DO TEMA NO CONTEXTO ATUAL

A interferência da Inteligência Artificial na vida humana se faz cada vez mais intensa e presente. Em virtude disso, é imprescindível discutir sobre o assunto, uma vez que ele está cada vez mais em voga no contexto dos avanços mercadológicos e sociais. Nesse cenário, nota-se o surgimento de novas relações sociais estabelecidas por sistemas e dispositivos interconectados, os quais passam a tomar decisões que antes eram consideradas exclusivamente humanas.

O assunto merece relevo, ainda, considerando principalmente a intensa atuação da Inteligência Artificial no setor privado, uma vez que é fatídica a tendência do mercado de incorporar o uso desse sistema para alavancar seu processo produtivo.

Nesse aspecto é um equívoco pensar que esse caminho se restringe à indústria ou a produção agrícola, uma vez que grandes empresas, bancos, e até operadores do Direito ao redor do mundo vêm se adaptando a essa realidade. Isso porque, na conjuntura globalizada atual, todas as ações voltadas a otimização do tempo para a realização do processo produtivo, são almejadas e perseguidas, independente do segmento.

O resultado lógico é a colaboração para que cada vez mais os sistemas de machine learning ocupem espaço na realidade mercadológica e social. Esse é um dos aspectos que denotam a importância da discussão ética e do controle dessa tendência por vias regulamentares, uma vez que os resultados podem ser altamente atentatórios aos Direitos Universais, contidos na Declaração Universal de Direitos Humanos de 1948.

Com base no contexto supracitado, na relevância do tema, nos dilemas éticos gerados e no ponto de contato dessa discussão com direitos tutelados, a Europa já tomou posição sobre o assunto. A partir de um grupo de especialistas ela lançou a “Ethics guidelines for trustworthy AI” (HIGH-LEVEL EXPERT GROUP ON AI, 2019b), um documento que visa apresentar princípios além de estipular exigências para que o desenvolvimento da Inteligência Artificial seja centrado no desenvolvimento humano e compatíveis com a realização de seus direitos. Tal documento será melhor analisado adiante

4. DIRETRIZES ÉTICAS DA UNIÃO EUROPEIA

Ante o exposto sobre a atualidade do tema, devido à complexidade e variedade de interpretações nos campos da moral e da ética, é necessário chegar a princípios basilares, mesmo que com alto grau de generalidade, para estabelecer parâmetros iniciais de análise.

Para tal, o desenvolvimento de IA que visem o bem comum em prol da humanidade deve, sem dúvida, considerar as necessidades humanas. Em virtude disso, com intuito de maximizar os benefícios da IA e, simultaneamente, prevenir e minimizar seus riscos, emerge a indispensabilidade, em um primeiro momento, de um sistema de normas que guiam o projeto de IA confiável, atendendo a todos os parâmetros citados abaixo.

Em 2018, a União Europeia lançou o primeiro panorama sobre as diretrizes a serem seguidas para a constituição da Inteligência Artificial. Após um *feedback* com diversos comentários, em 2019 o documento foi relançado pela comissão organizadora. As diretrizes se propõem à construção de uma IA confiável, e para isso esta deve ser lícita, ética e robusta, principalmente pelo fato de poderem ocorrer eventuais danos futuros, mesmo que não intencionais.²

O “Ethics guidelines for trustworthy AI” propõe uma definição do que seria, de fato, uma Inteligência Artificial de Confiança:

Extraímos a seguinte definição da literatura: Entende-se por confiança: 1) um conjunto de convicções específicas relacionadas com a benevolência, a competência, a integridade e a previsibilidade (convicções de confiança); 2) a disponibilidade de uma parte para depender

² Ethics guidelines for trustworthy AI, publicado em 8 de abril de 2019. Acesso em 20 de setembro de 2019. Disponível em https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419

de outros numa situação de risco (intenção de confiança); ou 3) a combinação destes elementos. Embora a confiança não seja uma propriedade atribuída a máquinas, o presente documento pretende salientar a importância de se poder confiar não só no facto de os sistemas de IA cumprirem a lei, respeitarem os princípios éticos e serem sólidos, mas também de se poder depositar tal confiança em todas as pessoas e todos os processos envolvidos no ciclo de vida do sistema de IA. (ETHICS GUIDELINES FOR TRUSTWORTHY AI, 2019b, p. 37)

Uma IA de confiança tem três componentes: 1) deve ser Legal, garantindo o respeito de toda a legislação e regulamentação aplicáveis; 2) deve ser Ética, demonstrando respeito e garantindo a observância de princípios e valores éticos; e 3) deve ser Sólida, tanto do ponto de vista técnico como do ponto de vista social, uma vez que, mesmo com boas intenções, os sistemas de IA podem causar danos não intencionais. Uma IA de confiança diz respeito não só à fiabilidade do próprio sistema de IA, mas também à fiabilidade de todos os processos e intervenientes que fazem parte do ciclo de vida do sistema. (ETHICS GUIDELINES FOR TRUSTWORTHY AI, 2019b, p. 37)

Logo, é oportuna a apropriação das orientações oferecidas pelo Guia, partindo da ideia de que a IA gera tanto benefícios quanto riscos para as sociedades, em geral. Principalmente quanto aos riscos relacionados às garantias do homem, seus direitos essenciais bem como nos efeitos do Estado Democrático de Direito. Nesse sentido, a premissa básica é que a IA precisa ser exclusivamente centrada no homem, e ser utilizada para maximizar os resultados dos serviços utilizados pela humanidade e promoção dos direitos humanos.

Dessa forma, torna-se necessário, inicialmente, compreender os quatro princípios norteadores das diretrizes que embasam a construção das exigências a serem seguidas na produção, na criação e na utilização direta da IA.

4.1 AUTONOMIA HUMANA

O primeiro princípio aborda o respeito pela autonomia humana, sendo imprescindível o tratamento digno das pessoas como sujeitos morais de direitos. Tal princípio visa reforçar a centralidade do homem e da sua integridade, física e mental, além da sua liberdade no processo de desenvolvimento das IA. Como a criação da Inteligência Artificial está voltada, exclusivamente, para o empoderamento humano, assim como das atividades realizadas por nós, aquelas devem ocupar um plano secundário de atuação, estando sob rígida vigilância humana. (ETHICS GUIDELINES FOR TRUSTWORTHY AI, 2019b, p.12)

4.2 PREVENÇÃO DE DANOS

O segundo princípio, da prevenção de danos, está intimamente relacionado à robustez mencionada anteriormente. Esta diferencia as categorias de danos possíveis de serem causados pelas IA, subdivididos naqueles inaceitáveis dos não intencionais. Para os primeiros devem haver grande esforço de prevenção e, para os segundos, a maior minimização possível. Por isso, esses sistemas devem estar aptos a fazer o melhor julgamento possível e evitar aprioristicamente os erros. Além disso, a robustez e a segurança também se relacionam com a capacidade de reprodutibilidade da IA averiguando se estas possuem o mesmo comportamento quando submetidas a mesma base de dados e as mesmas circunstâncias. (ETHICS GUIDELINES FOR TRUSTWORTHY AI, 2019b, p.12)

4.3 JUSTIÇA E EXPLICABILIDADE

O terceiro princípio, que disserta sobre a ética, busca equilibrar, de um lado, a justa e sensata distribuição dos custos e benefícios alcançados pela Inteligência Artificial de maneira igualitária, garantindo o tratamento paritário e, do outro, a ponderação entre os meios a serem seguidos e o fins alcançados.

No que diz respeito à seara procedimental, a justiça implica na possibilidade de discordância e reformulação das decisões tomadas pela IA ou pelas pessoas que as controlam. Por fim, a explicabilidade garante ser essencial com o intuito de que os usuários dos serviços produzidos por IA possuam confiança nos sistemas, os quais devem ser abertos e transparentes ao público, em geral. Estes princípios se encontram nos itens 52 e 53 do Ethics Guidelines for Trustworthy AI. (ETHICS GUIDELINES FOR TRUSTWORTHY AI, 2019b, p.12-13)

4.4 AS EXIGÊNCIAS

Os princípios acima mencionados se relacionam diretamente com as exigências impostas à criação e continuidade das IA, que devem possuir um sistema contínuo e ininterrupto de controle. As sete exigências citadas pelas diretrizes da UE são (i) *human agency* e supervisão humana, (ii) robustez técnica e segurança, (iii) privacidade e governança de dados, (iv) transparência, (v) diversidade, não discriminação e justiça, (vi) bem-estar ambiental e social e (vii) *accountability*. (ETHICS GUIDELINES FOR TRUSTWORTHY AI, 2019b, p.14)

4.4.1 HUMAN AGENCY E SUPERVISÃO HUMANA

A *human agency* e supervisão humana versa, claramente, sobre o controle humano em todas as etapas desenvolvidas da Inteligência Artificial (seja em cada etapa desenvolvida no ciclo, ou no conjunto final) para que esta, em nenhuma circunstância, possa atentar contra os direitos humanos já conquistados ou a forma democrática de governo. Dessa exigência também advém a necessidade de que os sistemas de Inteligência Artificial sejam de fácil manuseio ou compreensão, permitindo aos seus usuários o completo entendimento de qual procedimento está sendo realizado. (ETHICS GUIDELINES FOR TRUSTWORTHY AI, 2019b, p.15-16)

4.4.2 ROBUSTEZ TÉCNICA, SEGURANÇA E PRIVACIDADE

A exigência de robustez técnica e segurança está associada à necessidade de prevenção dos danos causados pela IA, sejam aqueles não intencionais, sejam aqueles inaceitáveis. Além disso, tal exigência tem por cabo afirmar a confiança e a reprodutibilidade dos sistemas, sendo essa a precisa repetição das decisões quando submetidas aos mesmos padrões.

No tocante à exigência de governança de dados e privacidade, os sistemas de IAs devem estar aptos para, antes mesmo de serem utilizados nos mercados, protegerem a base de dados da qual se abastecem, não permitindo o vazamento ou exposição destes. (ETHICS GUIDELINES FOR TRUSTWORTHY AI, 2019b, p.17)

4.4.3 TRANSPARÊNCIA

Ao mencionar a transparência, as exigências se propõem a deixar a Inteligência Artificial mais límpida e cristalina, tornando-se possível identificar porque esta tomou decisões de determinada

maneira, atingindo o cerne do problema. Dessa forma, transfiguram-se mais fácil explicar os processos técnicos e as decisões em si. Nesse ponto, é importante identificar o grau de influência que a IA exerce de maneira autônoma sobre as decisões.

Diretamente relacionada à transparência, estão a aplicabilidade e a comunicação. Esta última garante ao ser humano a veracidade dos fatos e das informações, sendo vedada a IA se apresentar como ser humano. (ETHICS GUIDELINES FOR TRUSTWORTHY AI, 2019b, p.18)

4.4.4 DIVERSIDADE, JUSTIÇA E NÃO DISCRIMINAÇÃO

A exigência de diversidade, justiça e não discriminação visa, além de garantir a inclusão, evitar que a IA seja induzida a semear preconceitos e injustiças através da influência humana ou dos maus modelos de governança, infelizmente existentes até hoje. As diretrizes também se preocupam com a presença da diversidade de opiniões a serem embutidas desde a fase inicial da IA, de modo a ressaltar a igualdade. Também se direciona a extirpar uma exploração desproporcional do consumidor ou modelos de concorrência desleal. (ETHICS GUIDELINES FOR TRUSTWORTHY AI, 2019b, p.19)

4.4.5 BEM-ESTAR SOCIAL E AMBIENTAL

A exigência do bem-estar social e ambiental frisa o comprometimento que deve existir tanto da parte da Inteligência Artificial, quanto de seus criadores e programadores para com o bem-estar da sociedade e do ambiente, frente à crescente preocupação com os ecossistemas, que passaram a constituir direitos de quarta geração essenciais. (ETHICS GUIDELINES FOR TRUSTWORTHY AI, 2019b, p.19)

4.4.6 ACCOUNTABILITY

Enfim, a exigência de *accountability*³, intrinsecamente relacionada à ética, concerne ao fato da averiguação do desenvolvimento da Inteligência Artificial e de observar se essa exerce suas funções, programações e objetivos que foram traçados e estabelecidos previamente (momento de sua criação).

Uma das preocupações emerge da possibilidade de os efeitos dessas evoluções tecnológicas impactarem diretamente direitos humanos, nesse caso, a exigência de transparência e fiscalização se tornarão ainda mais necessárias.

Dessa forma, o documento exige uma reavaliação intensa de cada fase processada e responsabiliza os desenvolvedores por cada impacto negativo gerado pela IA. Assim, essas reavaliações funcionam como uma espécie de ponderação ou *tradeoff* entre a existência ou não do sistema artificial (FRAZÃO, 2019). Portanto, se a implementação desse não puder seguir um padrão ético considerado, ele não pode existir.

³ O termo *accountability* pode ser entendido de diversas formas. No presente artigo, é utilizada a definição proposta pelo Guia da União Europeia, que visa possibilitar a avaliação dos algoritmos, dos dados e processos de concessão da inteligência artificial, o que não implica a total publicização das informações, mas sim a avaliação por auditores internos e externos e a disponibilidade desses relatórios, contribuindo para sua confiabilidade. Em situações que afetem direitos, o sistema de IA deve ser objeto de auditoria (*accountability*). (Ethics Guidelines for Trustworthy Artificial Intelligence. Disponível em https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419. Acesso em 20 de setembro de 2019.

Essas regulamentações fazem sentido à medida que se observa a ampla difusão dos sistemas de IA na vida social. Dessa forma, elas devem abarcar todos que, em algum grau, podem se relacionar com alguma Inteligência Artificial.

Principalmente diante das preocupações com igualdade, solidariedade e paridade, tantas vezes mencionadas nesse artigo, essas diretrizes se voltam à proteção, principalmente, dos grupos considerados vulneráveis como crianças, idosos, aqueles historicamente apartados das conquistas sociais, e aqueles que mantêm relações desiguais, como empresários e trabalhadores.

As diretrizes estabelecidas pela Comissão Europeia foram recebidas de maneira positiva no âmbito da Inteligência Artificial e na comunidade acadêmica. Por ser o primeiro grande movimento em prol da regulamentação da implantação de sistemas de Inteligência Artificial, essas diretrizes formam um padrão bem elaborado e já estão servindo de base para que outros países também pensem na regulamentação das IA.⁴

Trata-se de um assunto inadiável, visto o crescimento da atuação dessas tecnologias na vida da sociedade, em geral, e, majoritariamente, no setor privado.

Apesar de concretas e fundadas, no Direito sempre há espaço para a discussão, para a ampliação e para a melhoria de normas e regulamentações. (ETHICS GUIDELINES FOR TRUSTWORTHY AI, 2019b, p.19-20).

5. INTELIGÊNCIA ARTIFICIAL: CONTROLAR OU SER CONTROLADO

Muito antes das discussões sobre diretrizes éticas e possíveis parâmetros estabelecidos pela União Europeia, uma das controvérsias mais desafiadoras do desenvolvimento da IA era sua possível fonte e forma de regulamentação.

A *Microsoft*, ao lançar seu livro "*The Future Computed: Artificial Intelligence and its role in society*" (O Futuro Computadorizado: a Inteligência Artificial e seu papel na sociedade), retrata além das preocupações com o futuro da Inteligência Artificial, a necessidade de uma regulamentação específica e precisa. Assim, as empresas devem assumir responsabilidades frente aos sistemas inteligentes produzidos.

A empresa também delinea a preocupação com crises globais de desemprego que possam surgir ou possíveis agressões físicas e/ou verbais sofridas por humanos advindas dos robôs ou da IA. Diz a *Microsoft* em seu livro (MOREIRA, 2016):

A indústria da tecnologia precisa se engajar para mudar a percepção da mesma avançar em detrimento dos trabalhadores, que são substituídos sem proteção, benefícios, ou caminhos para suas carreiras a longo prazo. Empresas que não compreendem a importância da proteção e benefícios ao trabalhador podem prejudicar suas marcas e se surpreenderem com regulações futuras.

Assim, além das preocupações com a não segregação, não discriminação e com a igualdade, a IA gera, simultaneamente, benefícios e onerosidades tanto para os indivíduos diretamente relacionados com o seu desenvolvimento e uso (como grandes empresas), quanto para a sociedade em geral.

⁴ É o caso do Brasil, que no dia 27 de setembro de 2019 enviou para a relatoria do Senado Federal o projeto de lei nº 5051 de 2019, o qual contém 4 páginas e possui o objetivo de estipular uma regulamentação mínima sobre o uso de Inteligência Artificial no país.

Esses impactos negativos que podem advir são de difícil previsão, reconhecimento e determinação. Nesse sentido, a regulamentação tenta traçar formas de ponderar e garantir a existência e permanência do Estado Democrático de Direito, das garantias legais, dos direitos fundamentais do homem e até mesmo de sua existência e preponderância cognitiva. (DONEDA et al., 2018, p. 4)

Outro fator determinante que demonstra a necessidade de uma regulamentação, seria a existência de IA maléfica no tocante às relações sociais e jurídicas, como o caso das *deepfakes*. Essas são usadas para se referir a uma Inteligência Artificial que emprega técnicas de *deep learning* para produzir vídeos *fakes*.

A análise do caso é de extrema importância, pois revela como a escolha autônoma da IA é revestida de um dizer ético de previsão praticamente impossível e resultados inimagináveis. Não há controle dos dados que norteiam a decisão da IA, assim, a escolha realizada com base neles pode ser atentatória a direitos, além de causar mais insegurança no mundo digital, encurralando a ciência jurídica, a qual muito provavelmente não conseguirá apresentar soluções de modo eficaz sem regulamentação prévia.

Além disso, o Brasil historicamente resiste à regulamentação sobre temas que envolvam o mundo virtual, a exemplo do que aconteceu com os crimes cibernéticos, diferente de diversos outros países do mundo, somente em 2012⁵ o Brasil regulamentou e tipificou alguns crimes na *internet*, essa demora fez com que diversas condutas passassem ilesas mesmo ferindo gravemente direitos, assim compreender o tema e a necessária regulamentação pelo estudo do caso faz-se mister, para que não se cometa com a IA o mesmo equívoco de negligência.

Na prática, a produção de um vídeo *deepfake* se dá através da coleta de dados de um rosto humano, por meio de fotos e vídeos tradicionais existentes em um banco de dados de fácil acesso (KEHDY et al., 2018, p. 2). O mapeamento dos detalhes rosto humano se segue pela transposição desses dados, considerados pessoais e sensíveis, para uma imagem já existente, com o objetivo de inserir, de maneira indetectável ao olho humano, o indivíduo escolhido na situação desejada. Deste modo, uma pessoa, com acesso a um banco de dados e a alguns algoritmos, consegue simular situações falsas em vídeos e fotos, mesmo sem conhecimento técnico apurado. (GARATONNI, 2019)

Nesse sentido, a Universidade de Munique, na Alemanha criou uma Inteligência Artificial, nomeada como *XceptionNet*, que criou um banco de dados com *deepfakes* já existentes, o objetivo era que o sistema tivesse a capacidade de reconhecer a diferença entre um rosto real e um rosto simulado. Esse sistema foi criado com a finalidade de proteger eventuais danos à imagem e a privacidade, que são os principais afetados pelas *deepfakes* (GARATONNI, 2019).

Face ao que foi apresentado, não há dúvidas quanto à relevância legal e social das decisões tomadas com base em algoritmos e, em razão disso, surge a necessidade de analisar os impactos nos direitos humanos pautados nas diretrizes apresentadas.

⁵ A Lei Brasileira 12.737/2012, sancionada em 30 de novembro de 2012, promoveu alterações no Código Penal Brasileiro (Decreto-Lei 2.848 de 7 de dezembro de 1940), tipificando condutas que constituem os chamados delitos ou crimes informáticos.

6. IMPACTOS DA INTELIGÊNCIA ARTIFICIAL NOS DIREITOS HUMANOS

A Declaração Universal dos Direitos Humanos⁶ representou um marco na história da tutela de direitos que respeitam a dignidade e a individualidade humana bem como a repressão à tirania. É um documento elaborado por atores sociais das mais diversas tradições jurídicas e culturais, e representa a tutela universal de direitos, que foi incorporada em diversas Constituições Federativas pelo mundo. Essa Declaração foi proclamada em 1948, após a segunda guerra mundial, com o objetivo de ser uma norma que todas as nações mundiais persigam, fundamentada na paz, na justiça, na liberdade, no desenvolvimento, na dignidade e no valor da pessoa humana.

A Declaração contém trinta artigos que tutelam os direitos considerados imprescindíveis para o pleno gozo da vida em sociedade, por isso há a garantia à vida, à liberdade, à educação, ao trabalho, à igualdade, Entre outros. Nesse momento, resta importante também a tutela universal do direito à vida privada.

A partir desse preâmbulo, será vislumbrado - tendo por base os dilemas éticos, as diretrizes da União Europeia e o sonho regulamentar - como os sistemas de Inteligência Artificial se relacionam a esses direitos, com consequências, em boa parte, imprevisíveis, mas em outra, estampadas e sinalizadas.

Os artigos primeiro e décimo segundo da Declaração Universal dos Direitos Humanos serão apreciados e essa escolha baseou-se nos casos concretos, disponíveis na data de realização do presente estudo, que já apresentaram potencial lesivo de atingir, diretamente, os retromencionados direitos.

Ou seja, justifica-se a preocupação anteriormente levantada por meio de casos concretos os quais se relacionam com a utilização de Inteligência Artificial por empresas ao redor do mundo, principalmente nos países com alto índice de investimento no desenvolvimento de novas tecnologias. A seguir, expõem-se os direitos humanos universais apreciados no estudo (DECLARAÇÃO UNIVERSAL DE DIREITOS HUMANOS, 1948, p. 4-8):

I - Todos os seres humanos nascem *livres e iguais em dignidade e direitos*. São dotados de razão e consciência e devem agir em relação uns aos outros com espírito de fraternidade. (grifo nosso)

XII - *Ninguém será sujeito à interferência em sua vida privada*, em sua família, em seu lar ou em sua correspondência, nem a ataque à sua honra e reputação. Todo ser humano tem direito à proteção da lei contra tais interferências ou ataques. (grifo nosso)

Nos tópicos que se seguem os direitos humanos serão apreciados sob a ótica da atuação e desenvolvimento da IA na sociedade contemporânea. O que se almeja é promover uma reflexão de como a atuação da IA pode repercutir na tutela dos direitos humanos. E nesse ponto, algumas de suas características devem servir de premissas, afinal os direitos humanos além de fundados sobre o respeito pela dignidade e o valor de cada pessoa, portanto, entendidos de maneira uniforme, igualmente essencial a todos, são universais, inalienáveis, indivisíveis, inter-relacionados e interdependentes, sendo módico respeitar alguns direitos e desrespeitar outros.

⁶ Declaração Universal dos Direitos Humanos. Acesso em 20 de julho de 2019. Disponível em <https://nacoesunidas.org/direitoshumanos/declaracao/>

6.1 O CAMPO MINADO DA PRIVACIDADE

Desde a formação da humanidade existe a preocupação do homem em ter sua intimidade e vida protegidas. A era pós-moderna é caracterizada em peso pela tomada de espaços pela tecnologia na vida cotidiana, e, pois, a privacidade ganha novos contornos nos tempos atuais. Reconhecido na Declaração Universal dos Direitos Humanos, em seu art. XII, o direito à privacidade⁷ foi cuidadosamente inserido na tutela dos direitos humanos, revelando a sua importância para o indivíduo e o desenvolvimento social. No ordenamento jurídico brasileiro, esse direito faz parte do rol de direitos da personalidade, isto é, deve ser garantido e tutelado desde a primeira troca oxigênio-carbônica humana, ele está previsto no artigo 21 do Código Civil brasileiro de 2002.⁸

A proteção da privacidade, sendo assim, também pode ser entendida como uma forma de concretização do princípio da dignidade da pessoa humana. Sua violação, para fins desse estudo, faz-se presente na falta de controle e sanção do uso indevido e indiscriminado dos aparatos tecnológicos, que pode causar lesões irreparáveis a direitos humanos, ou seja, do mesmo modo que a sociedade pós-moderna cria inovações tecnológicas, também clama por novas formas de controle e proteção frente ao aumento dos riscos proporcionados por esses avanços, em virtude de seu uso descomedido e invasivo frente a uma baixa ou insuficiente regulamentação.

Considerando que a privacidade também tutela a honra, prevista expressamente na Declaração Universal de Direitos Humanos, e a imagem, direitos indisponíveis dos cidadãos, não se deve supor que apenas por utilizar as tecnologias os indivíduos tacitamente aceitam ser alvo de toda e qualquer espécie de intervenção, vista aqui como invasão de sua privacidade.

A título exemplificativo do impacto que o sistema de Inteligência Artificial pode trazer a esse direito, temos o caso *Alexa*, a *smart speaker* da *Amazon Echo* que vem gerando situações desconfortáveis a seus consumidores, por violações da privacidade.⁹

Situações como essas podem configurar grave violação à privacidade, que em relação ao espaço público e conforme o conceito ilustrado no estudo, significa o direito de não ter sua presença registrada, controlando as informações que lhe dizem respeito. Conforme exemplificado, não há mecanismo de controle em relação ao que pode ser armazenado ou a qual informação poderá ser apropriada indevidamente fruto de um problema de interpretação ou de outros problemas não mapeados.

Na sociedade que hoje é intitulada de Sociedade da Informação¹⁰, os mecanismos de captação de informações podem proporcionar uma imagem pessoal aprisionada para a posterioridade, visto que hoje a divulgação de informações pela internet atravessa barreiras espaciais e temporais. Dessa

⁷ Utiliza-se o entendimento de direito à privacidade preconizado por Stefano Rodotà, ou seja, como “o direito de manter o controle sobre suas próprias informações e de determinar a maneira de construir a sua própria esfera particular”. Assim, o cerne do direito à privacidade neste estudo é representado pela possibilidade de cada um controlar o uso das informações que lhe dizem respeito (RODOTÀ, 2008, p.15)

⁸ Art. 21. A vida privada da pessoa natural é inviolável, e o juiz, a requerimento do interessado, adotará as providências necessárias para impedir ou fazer cessar ato contrário a esta norma.

⁹ A *smart speaker* da *Amazon Echo*, denominada *Alexa*, passou a gravar conversas de seus consumidores e enviar as gravações sem as respectivas autorizações para outras pessoas, no caso específico, ocorreu a gravação de uma conversa entre um casal, a conversa foi enviada a um funcionário que comunicou ao casal a respeito da possibilidade de estar sendo vigiado pela *Amazon*. A empresa, que confirmou o erro, disse que não foi um problema de hack, mas sim um problema de interpretação equivocada por parte de *Alexa* durante a conversa do casal.

¹⁰ Conceito de Manuel Castells, que define a sociedade de informação como um conceito utilizado para descrever uma sociedade e uma economia que faz o melhor uso possível das tecnologias da informação e comunicação, no sentido de lidar com a informação, e que torna esta como elemento central de toda atividade humana (CASTELLS, 2000).

forma, a rigor, apenas o consentimento expresso pode afastar a ofensa à privacidade em certos espaços, como os espaços públicos, por exemplo. Percebe-se que nessa Sociedade da Informação o direito à privacidade recebe uma nova perspectiva, agora o referido direito seria visto como “o direito de controlar o uso que os outros fazem das informações que me dizem respeito”. (RODOTÀ, 2008, p.15).

Diante de todas as problemáticas que os avanços tecnológicos podem causar, é notável a necessidade de monitoramento e aprimoramento desses mecanismos, a fim de se evitar que elas provoquem mais malefícios que benefícios para quem deles usufrui.

Portanto, as exigências impostas pela comissão da União Europeia de especialistas no assunto mostram-se imprescindíveis, principalmente em relação à privacidade e a exigência da robustez técnica, visto que a prevenção de danos causados por sistemas de Inteligência Artificial não é algo que deva ser desconsiderado, tendo em vista que hoje esses sistemas são utilizados nos mais diversos ramos e pelas mais variadas empresas. Ou seja, o impacto dos danos de determinada IA pode até mesmo arruinar uma empresa ou disponibilizar dados privados sem autorização de diversas pessoas, ensejando indenizações exorbitantes pela violação à proteção dos dados e à privacidade dos sujeitos de direito.

É inadmissível a oferta sistemas de Inteligência Artificial para uso que não passam confiança em sua reprodutibilidade a seus usuários, e para haver a observância dessas diligências é necessário que esses sistemas sejam capazes de proteger sua base de dados. Por isso, é mister a observância desses parâmetros de reprodutibilidade e explicabilidade, uma vez que os erros considerados não-intencionais devem ser evitados e corrigidos, e os erros considerados inaceitáveis devem ser reparados imediatamente, pois não podem ser repetidos de forma alguma, principalmente nos casos que há envolvimento de direitos humanos, onde a ofensa a determinado direito pode ganhar um nível global. (FRAZÃO, 2019)

Além de ser um dos três princípios basilares da Declaração Universal dos Direitos Humanos, a autonomia humana também se faz presente nas diretrizes éticas propostas pela União Europeia (HIGH-LEVEL EXPERT GROUP ON AI, 2019b).

O princípio da privacidade demonstra, mais uma vez, a valorização da pessoa humana, portadora de valores éticos insuprimíveis, e seu papel central como ser autônomo, racional e formado de prerrogativas que nem mesmo o mais avançado dos sistemas consegue possuir e, dessa forma, tais sistemas devem atuar de modo secundário e sempre sob vigilância.

6.2 INTELIGÊNCIA ARTIFICIAL, INTERNET E DIGNIDADE, HÁ INTERSEÇÃO?

Como já apontado, a Inteligência Artificial é causa de alguns desafios que envolvem os direitos humanos, precipuamente entre aqueles que tutelam a dignidade diretamente. Isso porque, todos possuem esse escopo, afinal, no momento em que a Declaração Universal dos Direitos Humanos foi elaborada, sua maior preocupação era respaldar o homem em sua base mais essencial, uma vez que a Segunda Guerra Mundial mostrou o lado humano mais abominável ao promover perseguições de cunho religioso, cultural e sexual. (MOREIRA, 2016)

Superado seus grandes temores, o final do século XX foi marcado pelo surgimento de fortes Estados Sociais, que não só se preocupavam com uma igualdade formal, mas também com a

possibilidade de materializá-la, conscientes em abaixar índices de desigualdade e miserabilidade que rodeavam os países ao sul do globo.

Nesse contexto, o surgimento de grandes políticas sociais foi essencial para reduzir as desigualdades e concretizar garantias. Entre as quais podemos citar: o direito à vida, à liberdade, à segurança pessoal (art. III), a capacidade de gozar de direitos e liberdades sem qualquer diferenciação de raça, cor, sexo, religião, origem social e nacional (art. II) e o esforço e cooperação nacional para garantias econômicas, sociais e culturais, etc. (art. XXII) (DECLARAÇÃO UNIVERSAL DOS DIREITOS HUMANOS, 1948, p. 5-12)

Nesse diapasão, com o fortalecimento da internet e a dominação do uso de Inteligência Artificial com o intuito de melhorar a vida cotidiana e aprimorar os serviços prestados. Tais princípios, supracitados, precisam ganhar nova interpretação, incluindo a proteção dos indivíduos frente aos novos sistemas, uma vez que muitos não conhecem a dimensão e o grau de atingibilidade que podem advir da IA.

O uso de plataformas e redes sociais é uma realidade inegável. A maioria de nós utilizamos a internet para nos comunicar, encontrar parceiros, resolver problemas financeiros, produzir conteúdo e outros milhares de tipos de utilidades. Muitos fornecem dados de maneira simplista, sem saber qual dimensão aquelas informações podem atingir.

6.2.1 A DIGNIDADE DA PESSOA HUMANA E O CASO DO PERFIL *TAY*

Um exemplo considerável a ser destacado, nesse sentido, é o perfil *Tay*, lançado pela *Microsoft* em março de 2016 para atuar em redes sociais, afinal ele foi um dos primeiros perfis abertos que dialogou com uma ampla faixa etária e realizou diversos aprendizados autônomos em altíssima velocidade, ainda, esse episódio ensinou uma lição primordial no trato com a IA, a da necessidade e importância da análise de danos e de uma reação rápida quando engendrado quaisquer sinais de violações a direitos, principalmente à dignidade humana.

Além dos dados pré-estabelecidos, a Inteligência Artificial adquiria mais informações à medida que se relacionava com outros perfis nas redes, principalmente no *Twitter*. O público alvo buscado pela *Microsoft* foi de jovens entre 18 e 24 anos. Em menos de 24 horas no ar, a robô passou a emitir mensagens de cunho preconceituoso, racista, homofóbico e extremadas, como ataques ao presidente dos Estados Unidos, à época Barack Obama, e outras homenageando e comentando feitos históricos de Hitler. O perfil foi rapidamente desativado pela empresa, que mais tarde emitiu notas de justificção, acrescentando que a tecnologia precisava de reparos. (MOREIRA, 2016)

À época do lançamento do perfil pela *Microsoft*, as diretrizes estabelecidas pela União Europeia pairavam apenas no âmbito da cogitação. Hoje, em uma análise gradual do caso, é possível afirmar que o perfil *Tay* desrespeitou princípios como o respeito pela autonomia humana, violando diversas integridades mentais ao dispensar tratamentos dignos à sujeitos que são morais por direito. Como o desenvolvimento de IA se pauta no empoderamento humano, é necessário que os sistemas estejam sob rígida vigilância humana. (ETHICS GUIDELINES FOR TRUSTWORTHY AI, 2019b, p.12)

O produto da *Microsoft* também afrontou o princípio que trata sobre justiça, uma vez que benefícios e custos alcançados pelos sistemas devem ser distribuídos de maneira igualitária, garantindo tratamento paritário a todos. Por fim, ultrapassou os limites ao desconsiderar a

explicabilidade, uma vez que é inerente a este princípio que os usuários confiem nos sistemas de Inteligência Artificial. (ETHICS GUIDELINES FOR TRUSTWORTHY AI, 2019b, p.12-13)

No mesmo sentido, ao relacionar o caso exposto com a exigências estabelecidas pelo Guia para a elaboração de uma IA robusta, ética e confiável, claramente o controle humano nas etapas de construção do perfil foram relativizados, exemplificado por seus ataques de cunho racista e antissemita. (ETHICS GUIDELINES FOR TRUSTWORTHY AI, 2019b, p.15-16)

Em segundo lugar, foi violada a exigência sobre transparência, aplicabilidade e comunicação, uma vez que a *Microsoft* retirou a plataforma do ar sem qualquer explicação sobre quais problemas e, em qual etapa de desenvolvimento se encontravam esses problemas, fizeram com que a robô emitisse mensagens preconceituosas. (ETHICS GUIDELINES FOR TRUSTWORTHY AIL, 2019b, p.18)

No entanto, o efeito principal do perfil *Tay* foi violar a exigência sobre diversidade, justiça e não discriminação, que visa garantir a inclusão e evitar que a IA semeie preconceitos e injustiças por meio de influências humanas, se preocupando com a diversidade e igualdade. (ETHICS GUIDELINES FOR TRUSTWORTHY AI, 2019b, p.19)

Dessa forma, claramente foram violados princípios basilares para a construção de uma Inteligência Artificial que possa atuar de forma ética e responsável.

A *Microsoft* sequer foi questionada ou mesmo responsabilizada pelo lançamento da plataforma, já que na época as hipóteses de regulamentações e responsabilizações ainda eram escassas. (MOREIRA, 2016)

Esse é um importante exemplo de violação da dignidade de pessoas humanas, uma vez que o perfil foi atroz ao tocar em pontos tão sensíveis no discurso quanto à igualdade humana. No mesmo âmbito, é um excelente exemplo de como as diretrizes, mesmo como uma compilação basilar, já seriam úteis para evitar o surgimento de IA que viole direitos humanos.

7. CONSIDERAÇÕES FINAIS

O surgimento da nova Era de Inteligência Artificial, constituída sobre os alicerces da hiperconectividade, imprevisibilidade e insegurança, traz consigo a urgência de reflexões acerca do futuro global. Esse novo cenário tecnológico, gera benefícios incalculáveis para a desenvolvimento da humanidade, no entanto também ocasiona dilemas, que atualmente ainda são subestimados, os quais afetam diretamente os direitos humanos.

Verifica-se, no presente estudo, que o foco jurídico foi destinado a analisar alguns dos principais casos em que sistemas de Inteligência Artificial foram utilizados de maneiras desastrosas, ferindo gravemente determinados direitos humanos (privacidade e dignidade da pessoa humana). Como resultado, chegou-se à conclusão de que apesar das Diretrizes da União Europeia serem o início de uma regulamentação e possivelmente, no futuro, a base para uma regulamentação jurídica em outros países, essa ainda versa sobre diversos pontos de forma precária. Entre eles, está a falta de solução para não obstaculizar o desenvolvimento tecnológico e econômico em curso.

Sendo assim, é imprescindível o aperfeiçoamento das Diretrizes europeias e a criação de uma regulamentação brasileira que saiba sopesar, de forma cautelosa, benefícios e riscos, sob a ótica de garantia dos direitos humanos, com o intuito de mitigar os riscos de violações a estes direitos, bem como de oferecer uma estrutura sólida para o investimento em novas tecnologias que utilizem esses sistemas.

Ademais, fica claro que deve existir uma obediência a critérios rígidos em relação à atuação dos sistemas de Inteligência Artificial, a fim de se alcançar uma transparência em seus resultados, bem como a possibilidade de um monitoramento completo e eficaz. A não observância desses princípios poderá acarretar decisões questionáveis no âmbito moral e ético, especialmente aquelas que influenciarem de maneira significativa nos direitos humanos.

Por outro lado, deve ser avaliada a contraposição entre o alto nível de exigência e a rigidez técnica necessárias para que os sistemas de IA não fujam dos parâmetros éticos, porém, também não sirvam como desincentivo para pesquisa e inovação nessa área. Essa é a dificuldade encontrada pelos que tentam regulamentar o tema, uma vez que em virtude da difícil projeção estatística do nível de evolução desses sistemas, corre-se o risco da criação de uma legislação excessivamente abstrata (na tentativa de abranger todos os casos), ou de uma regulamentação insuficiente.

O discurso acerca da liberdade na tomada de decisões deve ser pautado pela aplicação de dados pessoais para substanciar decisões automatizadas, priorizando sempre a proteção dos direitos essenciais do homem. Assim a ética torna-se mecanismo eficiente para estruturar futuras regulamentações legislativas, se tornando base para que a Inteligência Artificial criada tenha como parâmetro primordial a proteção e o aprimoramento dos princípios fundamentais convencionados tanto no âmbito nacional quanto no internacional.

Por fim, urge a imprescindibilidade da elaboração de uma gestão democrática composta por grupos que detenham um conhecimento específico acerca do tema em conjunto com a sociedade civil, a fim de se estipular as finalidades da IA e sua regulação jurídica mediante os impactos acarretados no plano fático. Sem essa mobilização, a solução de vários dos problemas apontados pode se mostrar precipitada, casuística ou facilmente suscetível a convergir aos interesses econômicos.

Dessa forma, cabe, também, aos juristas um maior protagonismo no tocante a essas questões, com o intuito de alcançar propostas e instrumentos jurídicos que possam proporcionar a transformação das dificuldades apontadas em possibilidades de ações que contribuam para a valorização e evolução da vida humana.

REFERÊNCIAS BIBLIOGRÁFICAS | REFERENCES | REFERENCIAS

BRASIL. **Constituição da República Federativa do Brasil**. Promulgada em 5 de outubro de 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicaocompilado.htm.

CASTELLS, Manuel. **A era da informação: economia, sociedade e cultura**. In: A Sociedade em rede. São Paulo: Paz e Terra, 2000. v. 1.

DECLARAÇÃO UNIVERSAL DOS DIREITOS HUMANOS. Assembleia Geral das Nações Unidas em Paris. 10 dez. 1948.

DONEDA, D. C. M. et al. **Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal**. Pensar Revista de Ciências Jurídicas, v. 23, n. 4, p. 1 – 17, Outubro 2018.

DONEDA, D.C.M. **Considerações iniciais sobre banco de dados informatizados e o direito à privacidade**. In: TEPEDINO, Gustavo (org.) Problemas de direito civil-constitucional. Rio de Janeiro: Renovar, 2000, pp.111-136.

EPSTEIN, L; KING, G. **Pesquisa empírica em direito [livro eletrônico]: as regras de inferência**. São Paulo: Direito GV, 2013 (Coleção acadêmica livre) 7 Mb; Título original: The rules of inference. – Vários tradutores.

FOX, M. **This Immigration Attorney Is Using AI To Make Legal Access Affordable**. 2019. Disponível em: <https://www.forbes.com/sites/meimeifox/2019/03/15/this-immigration-attorney-uses-ai-make-legal-access-affordable/#3deb69fb5887>. Acesso em: 28/04/19.

FRAZÃO, A. **Quais devem ser os parâmetros éticos e jurídicos para a utilização da IA?** 2019. Disponível em: https://www.jota.info/paywall?redirect_to=//www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/quais-devem-ser-os-parametros-eticos-e-juridicos-para-a-utilizacao-da-ia-24042019. Acesso em: 27/04/19.

GARATTONI, B. **Cientistas criam primeira arma contra deepfakes**. Disponível em: <https://super.abril.com.br/tecnologia/cientistas-criam-primeira-arma-contradeepfakes/>. Acesso em: 10/05/19.

GORZONI, P. **Inteligência Artificial: Riscos para direitos humanos e possíveis ações**. 2019. Disponível em: <https://itsrio.org/wp-content/uploads/2019/03/Paula-Gorzoni.pdf>. Acesso em: 6 set. 2019.

HIGH-LEVEL EXPERT GROUP ON AI. **A definition of AI: Main capabilities and scientific disciplines**. Comissão Europeia, Bruxelas, p. 1 – 7, 2019a.

HIGH-LEVEL EXPERT GROUP ON AI. **Ethics Guidelines for Trustworthy AI**. Comissão Europeia, Bruxelas, p. 1 – 39, 2019b.

Homero. **Ilíada**. Tradução de Frederico Lourenço. Lisboa: Cotovia, 2005.

KAUFMAN, D. **Inteligência Artificial: questões éticas a serem enfrentadas**. In: ANAIS. 2016. IX Simpósio Nacional ABCiber. [S.l.]. 2016

KEHDY, B. O. et al. **Implicações Políticas, Sociais e Éticas da popularização de deepfakes por meio de softwares livres**. In: ANAIS, 2018. ANAIS do UEADSL. [S.l.], 2018.

MAGRANI, E. **Entre dados e robôs: ética e privacidade na era da hiperconectividade / Eduardo Magrani**. - 2. ed. - Porto Alegre: Arquipélago Editorial, 2019.

MOREIRA, I. **A Microsoft criou uma robô que interage nas redes sociais - e ela virou nazista**. 2016. Disponível em: <https://revistagalileu.globo.com/blogs/buzz/noticia/2016/03/microsoft-criou-uma-roboto-que-interage-nas-redes-sociais-e-ela-virou-nazista.html>. Acesso em: 28/04/19.

O'NEIL, C. **Weapons of math destruction: how big data increases inequality and threatens democracy**. Nova York: Crown Publishers, 2016.

PEREIRA, B. T.; SANTOS, S. R. do N.; GANDRA, G. G. Algoritmos e integração de novas tecnologias ao sistema jurídico. Disponível em: https://www.jota.info/paywall?redirect_to=//www.jota.info/opiniao-e-analise/colunas/regulacao-e-novas-tecnologias/algoritmos-e-integracao-de-novas-tecnologias-ao-sistema-juridico-27042019. Acesso em: 30/04/19.

RACANICCI, J. **Tecnologia introduz dilemas éticos ao Direito**. Disponível em: <https://www.jota.info/justica/tecnologia-introduz-dilemas-eticas-ao-direito-05122017>. Acesso em: 20/04/19.

RISSE, Matthias. **Direitos Humanos e Inteligência Artificial: Uma Agenda Urgentemente Necessária**. Revista Publicum, v. 4, n. 1, p. 17-33, 18 jun. 2018. Disponível em: <<http://www.e-publicacoes.uerj.br/index.php/publicum>>. Acesso em: 5 set. 2019.

RODOTÀ, Stefano. **A vida na sociedade de vigilância – A privacidade hoje**. Rio de Janeiro: Renovar, 2008. p. 15.

SILVEIRA, S. A. **Tudo sobre tod@s: Redes digitais, privacidade e venda de dados pessoais**. São Paulo: Edições SESC, 2017.

SMITH, Brad; SHUM, Harry. **O futuro computadorizado: a inteligência artificial e seu papel na sociedade**. Publicado por Microsoft Corporation Redmond, Washington. E.U.A. 2018.

SOUZA, C.A.; PADRÃO, V. **IA transformará o Direito, mas o Direito transformará IA?** Disponível em: <https://www.jota.info/opiniao-e-analise/artigos/ia-tranformara-o-direito-mas-o-direito-transformara-ia-26092017>. Acesso em: 20/04/19.

WAKKA, Wagner. In: **Alexa grava conversa privada de usuária e envia a gravação para outra pessoa**. [S. l.], 24 maio 2018. Disponível em: <https://canaltech.com.br/casa-conectada/alexa-grava-conversa-privada-de-usuaria-e-envia-a-gravacao-para-outra-pessoa-114512/>. Acesso em: 19 set. 2019.

YIN, Roberto K. **Estudo de Caso: planejamento e métodos**. 4ª ed. Tradução Ana Thorell. São Paulo: Bookman, 2010.

Ana Carolina de Souza Pereira

Graduanda do sexto período da Faculdade de Direito da UFJF.
<http://lattes.cnpq.br/9039881270937426>
anacarolinasp16@outlook.com

Ana Luisa Brêtas Bruno

Graduanda do sexto período da Faculdade de Direito da UFJF.
<http://lattes.cnpq.br/7372417361208120>
analuisa.bretas@gmail.com

Ariane Meirelles de Azevedo

Graduanda do sexto período da Faculdade de Direito da UFJF.
<http://lattes.cnpq.br/2246504332301872>
ariane.deazevedo@hotmail.com

Caroline da Rosa Pinheiro

Professora Adjunta do curso de Direito da Universidade Federal de Juiz de Fora (UFJF). Doutora em Direito pela UERJ, Mestre pela Fundação Getúlio Vargas - FGV/Direito – Rio.
<http://lattes.cnpq.br/5677415478719377>
caroline.ufjf@gmail.com

Livia Melo Honorato Campos

Graduanda do sexto período da Faculdade de Direito da UFJF.
<http://lattes.cnpq.br/8736622912082253>
livia.melo@direito.ufjf.br

Nicole Vieira Abreu Ornellas

Graduanda do sexto período da Faculdade de Direito da UFJF.
<http://lattes.cnpq.br/3850520464252435>
nicole.abreuorn@hotmail.com

Victória Presoti Paixão

Graduanda do sexto período da Faculdade de Direito da UFJF.
<http://lattes.cnpq.br/0222411460676641>
vicpaixao@yahoo.com.br