

Contribuição para a Comissão de Juristas Responsável por Subsidiar a Elaboração de Substitutivo Sobre Inteligência Artificial no Brasil

JOAO NETO <joaoneto@unifor.br>

sex 10/06/2022 17:49

Para: CJSUBIA <CJSUBIA@senado.leg.br>;

 1 anexo

GETIS - Contribuição SENADO_2022.pdf;

Você não costuma receber emails de joaoneto@unifor.br. [Saiba por que isso é importante](#)

Prezado Sr Presidente da Comissão de Juristas Responsável por Subsidiar a Elaboração de Substitutivo Sobre Inteligência Artificial no Brasil

Encaminhamos a V.Sia singela contribuição do Grupo de Estudos em Tecnologia, Informação e Sociedade da Universidade de Fortaleza.

Grato pela atenção,

João Araújo Monteiro Neto
Curso de Direito
Universidade de Fortaleza
Senior Reader in Law
Law School
University of Fortaleza

[Unifor.br](#) | [Instagram](#) | [Facebook](#) | [Twitter](#) | [LinkedIn](#) | [TV Unifor](#) | [G1/Ensinando e Aprendendo](#)



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

CONTRIBUIÇÃO

Para a elaboração de Marco Regulatório Sobre Inteligência Artificial

TIME DE PESQUISADORES

Ana Lourdes Teixeira Matos
André Câmara Ferreira da Costa
Francisco Marcelo de Souza Taumaturgo
Guilherme Fonsatti
Iago Capistrano Sá
Ikaró Rafael Miranda Fontenele
Isabelle Brito Bezerra Mendes
Letícia Magalhães Brito Alves
Narillyane Ferreira Guedes
William Magalhães Lessa
João Araújo Monteiro Neto

*Subsídio para o Marco
Normativo sobre Inteligência Artificial*

Fortaleza, Ceará
Maio de 2022



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

INTRODUÇÃO

A celeridade com que ocorrem transformações na atual sociedade é sem dúvida um tema intrigante que permeia as discussões e debates acadêmicos e sociais. A presença da tecnologia tem inclusive contribuído em grande parte para que essas alterações ocorram, trazendo mais comodidade ao cotidiano da humanidade e, por isso adentrando cada vez mais espaços da vida humana, inclusive os próximos a vida íntima.

A Inteligência Artificial ocupa esse lugar, cada vez mais os indivíduos têm optado por utilizar as máquinas para a celeridade de processos e atividades, para tomada de decisões e para realização dos mais diversos tipos de trabalhos. Isso tem trazido uma inovação enorme e uma possibilidade de qualidade de vida relevante para as pessoas. Não há mais como retroceder, é certo.

Entretanto, apesar das possibilidades advindas da presença da tecnologia na vida humana, especificamente a IA, por ser um campo relativamente novo que tem ganhado expansão nos últimos anos, não há como negar as questões negativas que se levantam nesse contexto. Por isso é preciso ser regulamentada, e aqui enfatizamos que não apenas por razões protetivas ou prevenção, mas também para assegurar ao seu usuário confiança, entendimento, acessibilidade e proteção.

Por isso, alguns países já têm se manifestado no sentido de normatizar o uso da IA, alguns até com diretrizes relativamente estruturadas para guiar de forma eficiente a vida prática. Outros países, como é o caso do Brasil, ainda estão buscando acompanhar essa caminhada normativa e precisam olhar para exemplos de outras nações para ter uma boa base para seus direcionamentos nesse sentido.

Portanto, é coerente que se busque inspirações em países que já têm caminhado nesse sentido, seja olhando para as medidas que se comprovaram eficazes, ou para outras que se mostraram infrutíferas. Nesse sentido, há diversos países bem encaminhados no que diz respeito a Inteligência Artificial e, para elucidar os eixos temáticos propostos, é



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

importante fazer o estudo de direito comparado para entender em que direção se tem caminhado.

Dentro disso, foi realizado levantamento de informações para contribuir com a tomada de subsídios para o Marco Normativo sobre Inteligência Artificial e de como tem sido a normatização dessa temática em países dos continentes América do Sul e do Norte e Europa. Ademais foram levados em consideração também doutrinas relevantes de autores que muito tem contribuído no estudo de Inteligência Artificial.

SUMÁRIO

INTRODUÇÃO

EIXO TEMÁTICO 1

1. CONCEITOS, COMPREENSÃO E CLASSIFICAÇÃO DE INTELIGÊNCIA ARTIFICIAL

1.1 Objeto a ser regulado

1.2. Aspectos sócio-técnicos

1.4 Princípios e objetivos

EIXO TEMÁTICO 3

3. DIREITOS E DEVERES

3.1 Transparência e Explicabilidade

3.3 Revisão e Direito à intervenção humana

3.7 Segredos Comercial e Industrial

CONCLUSÕES



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

EIXO TEMÁTICO 1

1. CONCEITOS, COMPREENSÃO E CLASSIFICAÇÃO DE INTELIGÊNCIA ARTIFICIAL

1.1 Objeto a ser regulado

Conceituar Inteligência Artificial (IA) é tarefa complexa no sentido de que há diferentes entendimentos do termo na literatura de referência, de acordo com enfoques seja da Ciência da Computação, da Engenharia do Direito e outros campos do conhecimento.

Peter Norvig e Steve Russel apresentam uma matriz de conceitos, dividindo-os de acordo com o destaque para *processos de conhecimento e raciocínio* (na parte superior) e de *comportamento semelhante ao humano* (na parte inferior). Confira-se¹:

Pensando como um humano	Pensando racionalmente
“O novo e interessante esforço para fazer os computadores pensarem (...) <i>máquinas com mentes</i> , no sentido total e literal.” (Haugeland, 1985) “[Automatização de] atividades que associamos ao pensamento humano, atividades como a tomada de decisões, a resolução de problemas, o aprendizado...” (Bellman, 1978)	“O estudo das faculdades mentais pelo uso de modelos computacionais.” (Charniak e McDermott, 1985) “O estudo das computações que tornam possível perceber, raciocinar e agir.” (Winston, 1992)
Agindo como seres humanos	Agindo racionalmente
“A arte de criar máquinas que executam funções que exigem inteligência quando executadas por pessoas.” (Kurzweil, 1990) “O estudo de como os computadores podem fazer tarefas que hoje são melhor desempenhadas pelas pessoas.” (Rich and Knight, 1991)	“Inteligência Computacional é o estudo do projeto de agentes inteligentes.” (Poole <i>et al.</i> , 1998) “AI... está relacionada a um desempenho inteligente de artefatos.” (Nilsson, 1998)

(Figura 1 – Matriz de conceitos de IA por Russel e Norvig)

Veja-se que os conceitos ou enfocam na possibilidade de a máquina reproduzir comportamentos semelhantes aos humanos, seja na possibilidade de ela replicar

¹ RUSSEL, Stuart; NORVIG, Peter. **Inteligência Artificial**. Tradução: Regina Célia Simille. 3ª ed. Edição Eletrônica. Rio de Janeiro, Elsevier, 2013, pos. 15.2.



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

processos semelhantes aos humanos, especialmente a capacidade de perceber o ambiente, processar tais percepções, tomar decisões e posteriormente agir sobre esse ambiente.

Na perspectiva de a IA *reproduzir* o comportamento humano, Alan Turing ainda em 1950, propôs um teste no qual um interrogador humano formularia determinadas perguntas por escritos para uma entidade, sem conhecer se o seu interlocutor seria um humano ou um computador. Na modalidade mais completa do teste, o robô precisaria, no mínimo, ter as seguintes capacidades: **a)** processamento de linguagem natural (PLN), para se comunicar com o interrogador humano em seu idioma; **b)** representação do conhecimento, para armazenar suas percepções; **c)** raciocínio automatizado, para usar o conhecimento armazenado para formular respostas e tirar conclusões; **d)** aprendizado de máquina, para se adaptar a novas circunstâncias e detectar e extrapolar padrões; **e)** visão computacional, para perceber objetos no ambiente; **f)** robótica, para manipular objetos e se movimentar.

Embora não seja de alta aplicação prática, o teste de Turing serviu para compilar as características necessárias para se atribuir a um computador o que seria efetivamente uma IA, sendo estes os principais campos de estudos de IA atualmente.

Stuart e Norvig comentam que a humanidade conseguiu desenvolver a tecnologia para voar quando deixou de replicar exatamente o que os pássaros faziam, mas a estudar a física e a aerodinâmica necessárias para o voo².

A autoria do termo IA em si é atribuída a John McCarthy, professor de Stanford que compilou um breve artigo com as principais perguntas sobre o tema. Deparado com a questão “O que é IA”, ele propôs a seguinte definição (tradução livre dos autores)³:

² RUSSEL, Stuart; NORVIG, Peter. *Ob. Cit.*, pos. 16.9

³ McCARTHY, John. **What is artificial intelligence?** Stanford University. Computer Science Department. Disponível em: < <http://jmc.stanford.edu/articles/whatisai/whatisai.pdf> >. Acesso em 10/05/2022. No original: It is the science and engineering of making intelligent machines, especially



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

IA é a ciência e a engenharia de fabricar máquinas inteligentes, especialmente programas de computador inteligentes. Está relacionada à tarefa similar de utilizar computadores para compreender a inteligência humana, mas a IA não precisa se limitar a métodos biologicamente observáveis.

Uma definição mais completa pode ser encontrada na literatura da Comissão Europeia, na figura do Grupo Independente de Peritos de Alto Nível sobre IA (GPAN IA) (tradução livre)⁴ que oferece uma definição mais atual para o termo IA:

IA se refere a sistemas produzidos por humanos que, dado um objetivo complexo, atuam no meio físico ou digital percebendo seu ambiente, interpretando os dados coletados de maneira estruturada ou não, raciocinando sobre o conhecimento derivado destes dados e decidindo acerca da melhor ação(ações) a tomar (de acordo com parâmetros previamente estabelecidos) para atingir o referido objetivo complexo.

Sistemas de IA pode ainda ser desenhados para aprender a adaptar seu comportamento mediante análise de como o ambiente é afetado por suas ações anteriores.

Como disciplina científica, IA inclui diferentes abordagens e técnicas, como aprendizagem de máquina (na qual aprendizagem profunda e aprendizagem por reforço são exemplos específicos), raciocínio de máquina (o qual inclui planejamento, agendamento, representação do conhecimento, além de raciocínio, pesquisa e otimização), e robótica (que inclui controle, percepção, sensores e atuadores, bem como a integração de todas as outras técnicas em sistemas ciber-físicos).

intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable.

⁴ COMISSÃO EUROPEIA. Grupo em Alto Nível de Experts em IA. **A definition of AI:** main capabilities and scientific disciplines. Bruxelas, dezembro de 2018. Disponível em: <https://ec.europa.eu/futurium/en/system/files/ged/ai_hleg_definition_of_ai_18_december_1.pdf>. Acesso em 11/05/2022. No original: “Artificial intelligence (AI) refers to systems designed by humans that, given a complex goal, act in the physical or digital world by perceiving their environment, interpreting the collected structured or unstructured data, reasoning on the knowledge derived from this data and deciding the best action(s) to take (according to pre-defined parameters) to achieve the given goal. AI systems can also be designed to learn to adapt their behaviour by analysing how the environment is affected by their previous actions. As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems).”



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

Portanto, nesta breve síntese de conceitos e definições, verifica-se que a IA hoje é, além de uma variedade de técnicas voltadas ao desenvolvimento de programas computacionais que replicam aspectos do comportamento ou do raciocínio humano, um campo do conhecimento próprio nascido de contribuições de diversas outras áreas do conhecimento.

No aspecto regulatório, é interessante destacar a importância dos estudos da Comissão Europeia para justamente delimitar sobre quais técnicas, aplicações e campos científicos uma normatização de IA incidiria – ponto de partida para se avaliar as consequências desta regulamentação.

Adiante, serão apresentadas algumas das diversas aplicações concretas de IA no cotidiano.

1.2. Aspectos sócio-técnicos

Conforme se detalhará no presente subtópico, as aplicações de IA atuais estão generalizadas tanto em serviços públicos quanto privados, cabendo, inicialmente, apresentar a compilação de prós e contras da aplicação de IA realizada por Fernanda Lage, partindo de estudos do Conselho de Políticas Públicas dos EUA (USACM), apresentadas em forma de tabela a seguir (adaptado pelos autores)⁵:

Vantagens e Desvantagens da Utilização de IA	
Vantagens	Desvantagens
- Diminuição da probabilidade de erro: ajuda a reduzir o erro e aumenta a chance de atingir a precisão com maior grau de exatidão. Descobrir erros, por exemplo, no início da produção de uma máquina é vital para que o processo seja o menos custoso possível. A IA pode ser utilizada para localizar qualquer alteração nos sistemas de	- Custos mais elevados: a criação de IA requer custos mais elevados, pois são máquinas complexas. Elas possuem programas que precisam de adaptação frequente para atender às necessidades do ambiente em constante mudança e a necessidade de as máquinas serem mais inteligentes a cada dia. No caso de avarias graves, o procedimento para recuperar códigos perdidos e

⁵ LAGE, Fernanda de Carvalho. **Manual de Inteligência Artificial no Direito Brasileiro**. Salvador: Editora JusPodvm, 2021, págs.47- 48.



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

<p>produção, de informação e de tecnologia, prevenindo e reduzindo possíveis erros existentes.</p> <ul style="list-style-type: none">- Utilização em ambientes hostis à vida e à saúde humana: a IA acoplada à robótica permite que trabalhos sejam feitos sem o risco da vida de pessoas;- Automatização de processos: automatizar processos é sinal de economia de tempo e de maior produtividade. A IA pode ser aplicada em processos que levam muito tempo e que são passíveis de erros quando feitos por um humano;- Gerenciamento e análise de dados e registros em áreas sensíveis: como, por exemplo, no campo financeiro, permitindo detectar anomalias e fraudes e organizar o capital da melhor maneira. É por isso que um mestre em IA é um dos recursos mais procurados pelos profissionais do setor.- Trabalhos repetitivos: trabalhos repetitivos de natureza monótona podem ser realizados com a ajuda da IA. As máquinas pensam mais rápidas que os humanos e podem ser multitarefas. A inteligência da máquina pode ser empregada para executar tarefas perigosas. Seus parâmetros, ao contrário dos humanos, podem ser ajustados, e sua velocidade e tempo são apenas parâmetros baseados em cálculo;- Melhoria da produtividade e da qualidade na produção: a IA não apenas aumenta a produtividade no nível de maquinaria, como também aumenta a produtividade do trabalhador e a qualidade do trabalho que realiza.	<p>restabelecer o sistema pode exigir muito tempo e custo.</p> <ul style="list-style-type: none">- Não possui criatividade original: os seres humanos são sensíveis e emocionais. Eles veem, ouvem, pensam e sentem. Seus pensamentos são guiados pelos sentimentos que faltam completamente nas máquinas. As habilidades intuitivas inerentes ao cérebro humano não podem ser replicadas.- Pode provocar o desemprego: a substituição de humanos por máquinas pode levar ao desemprego.
---	---

Embora na compilação realizada pela autora, os prós claramente ultrapassem os contras, no contexto das adaptações constantes necessárias para preservar o funcionamento da IA, é importante acrescer a possibilidade de comportamento



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

discriminatório da aplicação por falhas na programação, prejudicando determinado(s) grupo(s) de usuários.

Além disso, é preciso notar que uma decisão tomada por uma inteligência artificial não é autonomamente tomada por ela, mas anteriormente tomada pelo seu programador. Se compreendermos o papel do programador, que antecipa um esquema de solução, compreenderemos que o problema de decidir por um certo curso de ação é e continuará sendo humano. A diferença é que, na decisão sem a intervenção de uma inteligência artificial, o tomador de decisões é apenas o prestador direto dos serviços, enquanto na decisão com a intervenção participa também o programador, ainda que indiretamente. Isto significa que os problemas de decisão não se apagam com a simples contratação ou adoção de uma inteligência artificial para desempenhar o serviço, mas apenas se transferem para ou se partilham com um terceiro, que é o desenvolvedor da solução inteligente. E, na medida em que preservam, na raiz, uma dimensão humana, os problemas de decisão intercedidos por inteligências artificiais permanecem abertos, portanto, não menos difíceis nem mais acessíveis⁶.

A participação desse desenvolvedor-programador também não é sem dificuldades. Em primeiro lugar, é preciso ressaltar que o programador, ao ser contratado para elaborar uma solução computacional para uma demanda social, funciona como um tradutor das coordenadas dessa demanda, que surge em linguagem natural, para a linguagem lógico-matemática, algorítmica. No ínterim dessa tradução, podem ocorrer erros de interpretação, que são normais em toda relação humana, mas também falhas na adequação do produto à realidade do consumidor, que decorrem geralmente da ausência de uma formação multidisciplinar do desenvolvedor para entender as peculiaridades do mercado consumidor e, sobretudo, dos direitos em geral. Há, portanto, uma diferença

⁶ Nesse sentido, cf. LIMA, Sílvio Vieira Mendes. **Cérebros electrónicos e cérebros humanos: psíquico, psicológico e cibernético**. In: Revista Filosófica, 4, n. 2, 1952. BRONZE, Fernando Pinto. **Jurista: pessoa ou andróide?** In: VARELA et al. (Orgs.). AB VNO AD OMNES. 75 anos da Coimbra Editora. Coimbra: Coimbra Editora, 1998, p. 73-122



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

possivelmente problemática entre a lógica da computação e a lógica, por exemplo, do direito⁷.

Da parte do consumidor e, de modo geral, do jurisdicionado, a formação multidisciplinar, sobretudo em ciência da computação, também faz falta. Sem ela, não se tem condições de interpretar os resultados inesperados produzidos pela inteligência artificial. Acontece algo como um efeito “blackbox”, isto é, uma situação em que se conhecem tanto os dados informados ao aplicativo inteligente quanto os resultados que ele produziu, mas não se sabe o caminho que ele tomou para transformar os dados em resultados⁸. Isto dificulta o controle, a crítica e, conseqüentemente, a regulação dos problemas causados por esses dispositivos na vida social.

Outro ponto controverso na literatura é a questão de autoria e responsabilidade de processos conduzidos por IA, a exemplo dos carros autônomos e de obras artísticas elaboradas por IA – o que aumenta o desafio regulatório⁹.

Como veremos nos tópicos 3.1 e 3.3 deste trabalho, foi em atenção a essas desvantagens que os países desenvolveram alguns critérios para a utilização de inteligências artificiais nas atividades sociais. Esses critérios passam sempre pela necessidade de transparência quanto ao funcionamento dessas aplicações, com a possibilidade de revisão das decisões automatizadas, que devem ser explicáveis, além da exigência de uma intervenção humana mínima sobre elas.

Ademais, é importante ressaltar que os prós e contras apresentados podem variar muito de uma dada aplicação para outra, e o peso de uma determinada desvantagem pode

⁷ Nesse sentido, concluí já ARAÚJO, Fernando. **Lógica jurídica e informática jurídica** [1999]. In: Revista do Instituto do Direito Brasileiro. Ano 3 (2014), nº 2, p. 889-965.

⁸ BATHAEE, Yavar. The Artificial Intelligence Black Box and the Failure of Intent and Causation. In: Harvard Journal of Law & Technology, v. 31, n. 2, 2018.

⁹ Nesse sentido, vide a obra MEDON, Felipe. **IA e Responsabilidade Civil: autonomia, risco e solidariedade**. São Paulo: Editora Juspodivm, 2022.



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

se elevar e muito, diante, por exemplo, da adoção de IA em setores mais sensíveis, a exemplo da saúde.

Apresentadas as (des)vantagens da IA em geral, cabe exemplificar algumas de suas aplicações, conforme tabela adaptada a seguir¹⁰:

Setor	Aplicação de IA
Transportes	Sistemas detectam sinais de cansaço no rosto do motorista, outros assumem o controle completo do veículo e, outros ainda, tornam-se o cérebro do trânsito urbano, modelando a demanda por transporte público ou ajustando os semáforos para otimizar o fluxo de carros
Meio Ambiente	Redução da pegada de carbono e facilitar métodos de produção sustentáveis, desde a otimização de sistemas de filtragem de água até a regulação do consumo de energia de edifícios inteligentes ou promoção de agricultura frugal como insumo, estabelecimento de curto-circuitos ou a proteção de biodiversidade.
Indústria	Otimização de matérias-primas e estoques, manutenção preditiva ou logística inteligente em fábricas avançadas. A IA se tornou uma das características marcantes da chamada indústria 4.0, baseados em projetos que conectam soluções digitais, computação em nuvem, robótica e impressão 3D.
Saúde	A IA pode ser utilizada na relação paciente-médico, além de pesquisa, prevenção, diagnóstico e tratamento. O sistema Watson da IBM fornece assistência de diagnóstico e prescrição, principalmente com base no sequenciamento de DNA de tumores cancerígenos.
Serviços	Assistentes virtuais inteligentes podem ter um grande impacto no cotidiano das pessoas nos próximos anos. Conectados às plataformas de IA de seus designers, eles ativam ou desativam o aquecimento de sua casa conectada, gerenciam seu calendário, lembram-lhe de tomar seus remédios, encomendar suas tarefas, oferecer-se para ouvir música e fazer sua pesquisa sobre a rede.
Educação	A IA pode impactar sobre a educação em dois níveis: primeiro, o treinamento que os alunos devem receber para entendê-la e usá-la. Depois, há o

¹⁰ CHUI, Yadon *apud* LAGE, Fernanda de Carvalho. **Manual de Inteligência Artificial no Direito Brasileiro**. Salvador: Editora JusPodvm, 2021, págs.47- 48.



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

	conhecimento de que os seres humanos devem ter para viver em um mundo onde a IA irá se tornam onipresente.
--	--

Portanto, diante das diferentes áreas em que IA já é aplicada na sociedade, malgrado a inexistência de uma normatização específica para tais sistemas computacionais, os pesquisadores na área muito debatem não apenas sobre a necessidade de uma regulamentação e seus impactos sobre a inovação e a competitividade econômica do País, mas sobre a necessidade de uma regulamentação *generalizada* ou *setorizada*¹¹.

¹¹ Acerca deste debate, recomendam-se *podcast* do canal Dadocracia, do instituto Data Privacy Brasil, notadamente, episódios 78 e 80. Disponível em: <
<https://open.spotify.com/show/17Nu0CoPyQ0frpnQxyhLIo?si=a2ec9e6694274293>>.



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

1.4 Princípios e objetivos

Em busca de um mapeamento completo dos princípios de Inteligência Artificial, vale mencionar o levantamento realizado em conjunto pelo cewe.br, nic.br e cgi.br, na autoria de Caroline Burle e Diogo Cortiz. Nesse mapeamento, os autores levantam, a partir de uma análise dos princípios de I.A, seis dimensões, sendo elas: a Equidade (Fairness); Confiabilidade e Segurança (Reliability & Safety); Impacto Social (Social Impact); Responsabilidade (Accountability); Privacidade & Segurança (Privacy & Security); e Transparência (Transparency).

O mapeamento realizado pelos autores levou em consideração seis iniciativas internacionais, sendo duas do setor governamental (Comissão Europeia e Departamento de Defesa Norte-americano), duas do setor empresarial (Google e Microsoft), uma organização internacional (Organização para a Cooperação e Desenvolvimento Econômico – OCDE), e outra composta pela academia e setor empresarial (Academia de Inteligência Artificial de Pequim).

O intuito de escolha de tais iniciativas era garantir uma pluralidade regional, visto que não haviam encontrado nenhum material produzido para o Brasil. Tal mapeamento faz-se útil à contribuição, visto que levanta unanimidades e até mesmo distinções entre cada uma das iniciativas, podendo, desta forma, conduzir e determinar focos de pesquisas, ou ainda, determinar quais são as dimensões e conjuntos prático-filosóficos para se mapear os princípios de Inteligência Artificial e incluí-los, de maneira harmônica com o Ordenamento Jurídico Brasileiro, na legislação brasileira acerca das Inteligências Artificiais.



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

Princípios de IA. por dimensões	COMISSÃO EUROPEIA	OCDE	EMPRESAS E MAIS
Equidade	O desenvolvimento, a implantação e o uso de sistemas de Inteligência Artificial devem ser justos. Reconhecem que existem muitas interpretações diferentes de equidade (ou justiça), e dividem em duas dimensões: substantiva e procedural. A dimensão substantiva implica o compromisso de garantir uma distribuição equitativa e justa de benefícios e custos, assim como garantir que indivíduos e grupos estejam livres de preconceitos injustos, discriminação e estigmatização. Se desvios injustos puderem ser evitados, os sistemas de Inteligência Artificial podem até aumentar a justiça social.	Os sistemas de Inteligência Artificial devem ser projetados de maneira a respeitar o estado de direito, os direitos humanos, os valores democráticos e a diversidade, e devem incluir salvaguardas apropriadas - por exemplo, possibilitando a intervenção humana sempre que necessário - para garantir uma sociedade equitativa e justa.	Google: Evitar criar ou reforçar preconceitos injustos sobre as pessoas, particularmente aquelas relacionadas a características sensíveis, como raça, etnia, gênero, nacionalidade, renda, orientação sexual, habilidade e crenças políticas ou religiosas.
Confiabilidade e Segurança (está apenas como segurança no PL 21/2020)	Os sistemas de Inteligência Artificial não devem causar nem exacerbar danos ou afetar adversamente os seres humanos. Devem proteger a dignidade humana, bem como a integridade mental e física. Os sistemas de IA e os ambientes em que operam devem ser seguros, tecnicamente robustos e deve-se garantir que não estejam abertos ao uso malicioso. Pessoas vulneráveis devem receber maior atenção e ser incluídas no desenvolvimento, implantação e uso de sistemas de IA.	Os sistemas de Inteligência Artificial devem funcionar de maneira robusta, segura e protegida ao longo de seus ciclos de vida, e os riscos em potencial devem ser avaliados e gerenciados continuamente.	Google: Ser construído e testado para segurança para evitar resultados não intencionais que criam riscos de danos. Os sistemas de Inteligência Artificial serão projetados e desenvolvidos de acordo com as melhores práticas de pesquisa de segurança de IA. As tecnologias de IA serão testadas em ambientes restritos e a sua operação monitorada após a implantação.
Impacto Social (pode-se, por analogia, entender como o princípio de “Centralidade no ser humano” no PL 21/2020)	Os direitos fundamentais em que se baseia a União Europeia visam garantir o respeito pela liberdade e autonomia dos seres humanos. Os seres humanos que interagem com os sistemas de IA devem ser capazes de manter a autodeterminação completa e eficaz sobre si mesmos e participar do processo democrático. Os sistemas de IA não devem injustificadamente subordinar, coagir, enganar, manipular, condicionar ou agrupar humanos. Em vez disso, eles devem ser projetados para aumentar, complementar e capacitar as habilidades cognitivas, sociais e culturais humanas. A alocação de funções entre humanos e sistemas de IA deve seguir os princípios de design centrado no ser humano (human centered design) e deixar oportunidades significativas para a escolha humana	A Inteligência Artificial deve beneficiar as pessoas e o planeta, impulsionando o crescimento inclusivo, o desenvolvimento sustentável e o bem-estar.	Google: Ser socialmente benéfico e manter altos padrões de excelência científica. Levar em conta uma ampla gama de fatores sociais e econômicos e prosseguir onde acreditam que os benefícios prováveis gerais excedam substancialmente os riscos e desvantagens previsíveis. Disponibilizar informações precisas e de alta qualidade usando a IA, de forma a respeitar as normas culturais, sociais e legais nos países em que opera.



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

Responsabilidade	<p>Inclui auditoria, minimização e relatórios de impacto negativo e trade-offs. O requisito de responsabilidade está intimamente ligado ao princípio da equidade. É necessário que sejam criados mecanismos para garantir a responsabilidade e a prestação de contas dos sistemas de IA e seus resultados, antes e depois de seu desenvolvimento, implantação e uso.</p>	<p>As organizações e indivíduos que desenvolvem, implantam ou operam sistemas de Inteligência Artificial devem ser responsabilizados pelo seu bom funcionamento, de acordo com os princípios acima.</p>	<p>Departamento de Defesa Norte-americano: Os seres humanos devem exercer níveis adequados de julgamento e permanecer responsáveis pelo desenvolvimento, implantação, uso e resultados dos sistemas de Inteligência Artificial do Departamento de Defesa. Google: Ser responsável perante as pessoas. Projetar sistemas de Inteligência Artificial que ofereçam oportunidades apropriadas para feedback, explicações relevantes e apelo. As tecnologias de IA estarão sujeitas à direção e controle humanos adequados.</p>
Privacidade & Segurança	<p>Inclui respeito pela qualidade da privacidade e integridade dos dados e acesso aos dados. Intimamente ligada ao princípio de prevenção de danos está a privacidade, um direito fundamental particularmente afetado pelos sistemas de IA. A prevenção de danos à privacidade também exige governança de dados adequada que cubra a qualidade e a integridade dos dados utilizados, sua relevância à luz do domínio em que os sistemas de IA serão implantados, seus protocolos de acesso e a capacidade de processar dados de maneira que proteja a privacidade.</p>		<p>Google: Incorporar princípios de privacidade desde a concepção do projeto (privacy by design) . Incorporar princípios de privacidade no desenvolvimento e uso de nossas tecnologias de IA. Dar oportunidade de aviso e consentimento, incentivar arquiteturas com salvaguardas de privacidade e fornecer transparência e controle adequados sobre o uso de dados</p>
Transparência	<p>É crucial para criar e manter a confiança dos usuários nos sistemas de Inteligência Artificial. Isso significa que os processos precisam ser transparentes, as capacidades e o objetivo dos sistemas de IA comunicados abertamente e as decisões - na medida do possível - explicáveis para os afetados direta e indiretamente. Sem essas informações, uma decisão não pode ser devidamente contestada. Uma explicação sobre por que um modelo gerou uma saída ou decisão específica (e que combinação de fatores de entrada contribuiu para isso) nem sempre é possível.</p>	<p>Deve haver transparência e divulgação responsável em torno dos sistemas de Inteligência Artificial para garantir que as pessoas entendam os resultados baseados em IA e possam desafiá-los.</p>	<p>Google: Fornecer transparência apropriada e controle do uso de dados. Este princípio é citado dentro de Privacidade & Segurança. Academia de Inteligência Artificial de Pequim: A Pesquisa e Desenvolvimento da Inteligência Artificial deve adotar abordagens éticas de design para tornar o sistema confiável. Isso pode incluir, mas não se limita a: tornar o sistema o mais justo possível, reduzir possíveis discriminações e preconceitos, melhorar sua transparência, prover explicação e previsibilidade e tornar o sistema mais rastreável, auditável e responsável.</p>



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

EIXO TEMÁTICO 3

3. DIREITOS E DEVERES

3.1 Transparência e Explicabilidade

Para essa subdivisão especificamente foram analisados as instituições e países a seguir dispostos:

- Organização para a Cooperação e Desenvolvimento Econômico (OCDE), que foi a primeira instituição a trazer uma norma internacional relativa ao uso de IA, que serve de base para formalização de leis e diretrizes de diversos países. Trata-se portanto de Recomendações emitidas por seu Conselho de Inteligência Artificial.

(Documento disponível em: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>)

- Grupo Independente de Peritos de Alto Nível sobre Inteligência Artificial (GEPAN) que foi criado pela Comissão Europeia em 2018, sendo composto por 52 peritos independentes em representação do mundo acadêmico, da indústria e da sociedade civil, eles foram responsáveis por fornecer conselhos sobre a Estratégia de Inteligência Artificial europeia. O grupo possui uma série de entregáveis nesse contexto que muito podem auxiliar no entendimento da normatização da IA. Dentre os documentos estão Diretrizes Éticas para IA Confiável e Uma definição de IA: Principais Capacidade e Disciplinas Científicas.

(Documentos disponíveis em: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>)

- Comitê Ad hoc de Inteligência Artificial (CAHAI) que hoje é o Comitê de Inteligencia Artificial (CAI) do Conselho Europeu, que foi responsável por examinar a viabilidade e os elementos potenciais, com base em amplas consultas multissetoriais, de um quadro jurídico para o desenvolvimento, concessão e aplicação da inteligência artificial, com base nas normas do Conselho da Europa



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

sobre direitos humanos, democracia e Estado de direito. Em 2020 esse comitê realizou um estudo sobre a regulamentação dos sistemas de IA e publicou o documento Towards Regulation of AI systems.

(Documento disponível em: <https://www.coe.int/en/web/artificial-intelligence/-/-toward-regulation-of-ai-systems-#:~:text=The%20publication%20of%20the%20CAHAI,is%20nowadays%20of%20outstanding%20importance.>)

- Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO), com sua primeira recomendação sobre ética da inteligência artificial, que fixa como princípios a transparência de dados e a preservação do ambiente, tendo seu texto aprovado por 193 membros.

(Documento disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000381137>)

- Chile que possui uma Política de Inteligência Artificial que segue inclusive boa parte das orientações da OCDE

(Documento disponível em: https://minciencia.gob.cl/uploads/filer_public/bc/38/bc389daf-4514-4306-867c-760ae7686e2c/documento_politica_ia_digital_.pdf)

- Uruguai que possui uma Estratégia de Inteligência Artificial

(Documento disponível em: <https://www.gub.uy/agencia-gobierno-electronico-sociedad-informacion-conocimiento/sites/agencia-gobierno-electronico-sociedad-informacion-conocimiento/files/documentos/publicaciones/Estrategia%20IA%20-%20versi%C3%B3n%20espa%C3%B1ol.pdf>.)

- Colômbia que possui uma Política Nacional para a Transformação Digital e IA publicada em 2019, e outro documento publicado em 2020 intitulado Ethical Framework for Artificial Intelligence in Colombia vinculado a presidência da república.



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

CHILE

Em sua Política de Inteligência Artificial, o Chile também traz em seu texto os exatos 5 princípios recomendados pela OCDE, ademais ao falar sobre IA inclusiva traz a seguinte disposição:

Devido à estreita dependência que existe entre o treinamento de algoritmos e o uso de dados para isso, transparência e explicabilidade tornam-se elementos relevantes para a concepção de uma IA inclusiva, portanto, as ações darão especial ênfase aos atributos de integridade e qualidade dos dados para garantir que seus preconceitos sejam conhecidos e tratados adequadamente.

A IA não deve discriminar com base em categorias protegidas ou ser usada em detrimento de qualquer grupo. Em particular, é de especial importância que a IA seja desenvolvida com uma perspectiva de gênero e diversidade sexual, incluindo grupos historicamente relegados, como povos indígenas, pessoas com habilidades especiais ou os setores mais vulneráveis de nossa economia, para se tornar um instrumento útil para as pessoas, garantindo a redução e fechamento de lacunas.

Quanto a explicabilidade define da seguinte forma:

Explicabilidade: Este termo refere-se a métodos e técnicas que são utilizados para que a inteligência artificial e seus resultados possam ser explicados, ou melhor, compreendidos por especialistas.

URUGUAI

O Uruguai, em Estratégia de Inteligência Artificial, também traz no início de seu texto 6 princípios, os quais são:

1. Finalidade;
2. Interesse Geral;
3. Respeito aos Direitos Humanos;
4. Transparência;
5. Responsabilidade;
6. Ética.



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

No que se refere a transparência traz o seguinte texto:

Transparência: as soluções de IA utilizadas na esfera pública devem ser transparentes, em conformidade com as normas vigentes. Essa transparência deve:

- Disponibilizar os algoritmos e dados utilizados para o treinamento da solução e sua implementação, bem como os testes e validações realizados.
- Tornar explicitamente visível, por meio de mecanismos ativos de transparência, todos os processos que utilizam a IA, seja na geração de serviços públicos ou no apoio à tomada de decisões.

COLÔMBIA

A Colômbia, em sua Política Nacional para a Transformação Digital e IA, expõe um rol de 10 princípios, dentre eles há a seguinte disposição sobre transparência:

Marco ético para IA e segurança: o governo nacional reconhece que o uso de IA envolve uma série de desafios éticos que devem ser considerados e abordados pelo Estado, como justiça, liberdade, não discriminação, transparência, design responsável, segurança, privacidade e o papel dos direitos humanos, entre muitos outros. Esses princípios devem ser discutidos e construídos com o apoio do setor privado e da comunidade científica e acadêmica do país, especialistas na área.

Já em relação a explicabilidade diz:

Este princípio é visto às vezes como complementar à transparência, mas difere na medida em que as informações que são compartilhadas e divulgadas ao público de forma transparente devem ser compreensíveis tanto para os desenvolvedores e usuários do sistema quanto para aqueles indivíduos que podem ser afetados por suas decisões e resultados. Portanto, vai além da transparência e busca tornar o conteúdo da informação e a forma como ela é apresentada facilmente acessível, considerando as complexidades que cercam essa tecnologia.



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

Na ética dos dados: a explicação permite que as pessoas compreendam a importância dos dados na concepção e desenvolvimento desses sistemas, como coletam e processam informações e as finalidades para as quais o fazem, especialmente no processamento de dados pessoais.

Na ética dos algoritmos: significa que os algoritmos podem ser compreendidos no sentido de que se permite compreender os objetivos iniciais que se busca e que são próprios do modelo, e os resultados esperados e também os obtidos. É claro que nos sistemas conhecidos como “Caixa Preta” esse princípio pode entrar em tensão, uma vez que a capacidade de explicação pode ser limitada. Mesmo os desenvolvedores desse tipo de tecnologia e especialistas na área não conseguem entender completamente o processamento que ocorre e a maneira como os sistemas atingem determinados resultados. No entanto, pretende-se também explicar de forma compreensível como funcionam estes sistemas, porque são denominados Black Box e as implicações que isso pode ter. Da mesma forma, e em qualquer caso, devem ser apresentadas informações compreensíveis e claras sobre os objetivos perseguidos por este sistema no seu desenvolvimento e implementação.

Na ética da prática: significa fornecer informações claras e precisas sobre os papéis das pessoas envolvidas no projeto, desenvolvimento e implementação dessa tecnologia. Além disso, informações claras, precisas e compreensíveis devem ser geradas ao longo do processo sobre como esses sistemas estão sendo avaliados e mecanismos específicos devem ser gerados para compartilhar informações sobre os resultados obtidos, principalmente com as comunidades que estão sendo impactadas por esses sistemas.

No documento intitulado Ethical Framework for Artificial Intelligence in Colombia, traz mais algumas disposições sobre transparência, inclusive mais específicas:

A transparência deve ser entendida como abertura para fornecer informações completas sobre o projeto, operação e impacto dos sistemas de inteligência artificial. Tais informações não devem comprometer a confidencialidade do modelo de negócios e inovação, tornando-as suscetíveis ao uso não autorizado.

Na ética dos dados: a transparência tem o efeito de fornecer informações sobre as fontes das informações utilizadas para o design desta tecnologia, suas características e as finalidades para as quais os dados, especialmente os dados pessoais, serão usados. Isso afeta os dados utilizados para treinamento e



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

implementação desses sistemas, pois os critérios utilizados para sua classificação e processamento devem ser transparentes em sua coleta.

Na ética dos algoritmos: significa transparência no modelo por trás dos algoritmos, informando os critérios que levam esses sistemas a terem determinados tipos de resultados. Neste ponto, é relevante fornecer informações aos cidadãos sobre os insumos utilizados no desenho dos sistemas e os resultados que podem ser apresentados (transparência dos insumos e dos produtos). Além disso, informações completas devem ser fornecidas sobre os falsos positivos ou negativos que um algoritmo desenvolveu e as porcentagens de precisão (precisão do algoritmo).

Na ética das práticas: consiste em fornecer informações completas sobre as equipes e pessoas envolvidas no desenho e desenvolvimento desses sistemas, os objetivos que perseguem e os manuais de conduta que geraram para realizar seu trabalho. Esses manuais devem ser acessíveis ao público. Como prática para os responsáveis pela implementação desta tecnologia, deve ser privilegiado o uso de sistemas de código aberto, especialmente nos entes do Estado.

**GRUPO INDEPENDENTE DE PERITOS DE ALTO NÍVEL SOBRE
INTELIGÊNCIA ARTIFICIAL CRIADO PELA COMISSÃO EUROPEIA EM
2018**

O GPAN IA em suas Diretrizes Éticas para IA Confiável que trazem as seguintes considerações sobre transparência:

Este requisito está estreitamente relacionado com o princípio da explicabilidade e abrange a transparência dos elementos relevantes para um sistema de IA: os dados, o sistema e os modelos de negócio.

Rastreabilidade. Os conjuntos de dados e os processos que produzem a decisão do sistema de IA, incluindo os processos de recolha e etiquetagem dos dados, bem como os algoritmos utilizados, devem ser documentados da melhor forma possível para permitir a rastreabilidade e um aumento da transparência. Isto também se aplica às decisões tomadas pelo sistema de IA. Deste modo, é possível identificar os motivos por que uma decisão de IA foi errada, o que, por sua vez, poderá ajudar a evitar erros futuros. A rastreabilidade facilita, assim, a auditabilidade e a explicabilidade.



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

Explicabilidade. A explicabilidade diz respeito à capacidade de explicar tanto os processos técnicos de um sistema de IA como as decisões humanas com eles relacionadas (p. ex., os domínios de aplicação de um sistema de IA). A explicabilidade técnica exige que as decisões tomadas por um sistema de IA possam ser compreendidas e rastreadas por seres humanos. Além disso, poderá ser necessário adotar soluções de compromissos entre o reforço da explicabilidade de um sistema (o que poderá reduzir a sua exatidão) ou o aumento da sua exatidão (à custa da sua explicabilidade). Sempre que um sistema de IA tenha um impacto significativo na vida das pessoas, deverá ser possível solicitar uma explicação adequada do respetivo processo de tomada de decisões. Tal explicação deve ser oportuna e adaptada ao nível de especialização da parte interessada em causa (p. ex., leigo, regulador ou investigador). Além disso, devem ser disponibilizadas explicações sobre o grau de influência e de intervenção de um sistema de IA no processo decisório da organização, as opções de conceção do sistema e os fundamentos da sua implantação (assegurando assim a transparência do modelo de negócio).

Comunicação. Os sistemas da IA não se devem apresentar como seres humanos aos utilizadores; os seres humanos têm direito a serem informados de que estão a interagir com um sistema de IA. Tal implica que os sistemas de IA devem ser identificáveis como tal. Além disso, deve ser facultada a opção de decidir contra essa interação a favor da interação humana, sempre que necessário, a fim de garantir que os direitos fundamentais são respeitados. Além disso, as capacidades e limitações do sistema de IA devem ser comunicadas aos profissionais no domínio da IA ou aos utilizadores finais de forma adequada ao caso de utilização em questão. Essa comunicação poderá incluir o nível de exatidão do sistema de IA, bem como as suas limitações.

Já em relação a explicabilidade assim dispõem:

A explicabilidade diz respeito à capacidade de explicar tanto os processos técnicos de um sistema de IA como as decisões humanas com eles relacionadas (p. ex., os domínios de aplicação de um sistema de IA). A explicabilidade técnica exige que as decisões tomadas por um sistema de IA possam ser compreendidas e rastreadas por seres humanos. Além disso, poderá ser necessário adotar soluções de compromissos entre o reforço da explicabilidade de um sistema (o que poderá reduzir a sua exatidão) ou o aumento da sua exatidão (à custa da sua explicabilidade). Sempre que um sistema de IA tenha um impacto significativo na vida das pessoas, deverá ser possível solicitar uma explicação adequada do respetivo processo de tomada de decisões. Tal explicação deve ser oportuna e adaptada ao nível de especialização da parte interessada em causa (p. ex., leigo, regulador ou investigador). Além disso,



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

devem ser disponibilizadas explicações sobre o grau de influência e de intervenção de um sistema de IA no processo decisório da organização, as opções de concepção do sistema e os fundamentos da sua implantação (assegurando assim a transparência do modelo de negócio).

A explicabilidade é crucial para criar e manter a confiança dos utilizadores nos sistemas de IA. Tal significa que os processos têm de ser transparentes, as capacidades e a finalidade dos sistemas de IA abertamente comunicadas e as decisões — tanto quanto possível — explicáveis aos que são por elas afetados de forma direta e indireta. Sem essas informações, não é possível contestar devidamente uma decisão. Nem sempre é possível explicar por que razão um modelo gerou determinado resultado ou decisão (e que combinação de fatores de entrada contribuiu para esse efeito). Estes casos são designados por algoritmos de «caixa negra» e exigem especial atenção. Nessas circunstâncias, podem ser necessárias outras medidas da explicabilidade (p. ex., a rastreabilidade, a auditabilidade e a comunicação transparente sobre as capacidades do sistema), desde que o sistema, no seu conjunto, respeite os direitos fundamentais. O grau de necessidade da explicabilidade depende em grande medida do contexto e da gravidade das consequências de um resultado errado ou inexacto

OCDE

Vale ressaltar que logo no início de seu texto de recomendações a OCDE já sujere a implementação de 5 princípios são eles:

1. Crescimento inclusivo, desenvolvimento sustentável e bem-estar;
2. Valores centrados no ser humano e justiça;
3. Transparência e explicabilidade;
4. Robustez, segurança e proteção;
5. Prestação de contas.

No que se refere a transparência e a explicabilidade tem disposto o seguinte:

Os atores de IA devem se comprometer com a transparência e divulgação responsável em relação aos sistemas de IA. Para tanto, devem fornecer informações significativas, adequadas ao contexto e consistentes com o estado da arte:

- i. para promover uma compreensão geral dos sistemas de IA,



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

- ii. conscientizar as partes interessadas sobre suas interações com sistemas de IA, inclusive no local de trabalho,
- iii. para permitir que os afetados por um sistema de IA entendam o resultado e,
- iv. para permitir que aqueles afetados negativamente por um sistema de IA contestem seu resultado com base em informações claras e fáceis de entender sobre os fatores e a lógica que serviu de base para a previsão, recomendação ou decisão.

CONSELHO EUROPEU

Transparência: Apresentada em 101 das 116 fontes, a transparência é o princípio ético mais prevalente no atual espectro de soft law. A análise temática revela uma variação significativa em relação à interpretação e justificação dos apelos à transparência. Observa-se que essa variação causa divergências óbvias nas estratégias de implementação propostas para obter transparência em relação à IA.

As referências à transparência podem ser agrupadas em duas famílias temáticas principais: (1) transparência de algoritmos e métodos de processamento de dados, (2) transparência de práticas humanas relacionadas ao projeto, desenvolvimento e implantação de sistemas de IA.

Os apelos à transparência do tipo 1 normalmente envolvem a promoção de abordagens metodológicas para "IA explicável", ou seja, sistemas de IA cujas saídas e decisões podem ser compreendidas por especialistas humanos. Esses métodos e técnicas contrastam com as abordagens de "caixa preta" para aprendizado de máquina, onde as etapas pelas quais um sistema de IA chegou a uma decisão específica são ininteligíveis para especialistas humanos, incluindo os designers do sistema. Enquanto as empresas privadas, especialmente os atores privados de IA, tendem a reduzir a transparência à interpretabilidade e explicabilidade por meio de soluções técnicas – como, entre outras, propagação de relevância em camadas (LRP) e interpretabilidade local – órgãos governamentais, como oficiais nacionais de proteção de dados, enfatizam a importância da supervisão métodos como auditorias. Os apelos à transparência do tipo 2 não se concentram em algoritmos interpretáveis, mas na transparência de práticas humanas relacionadas a dados e IA, como divulgar informações relevantes aos titulares de dados, evitar sigilo ao implantar estratégias de IA e proibir conflitos de interesse entre atores de IA e órgãos de supervisão . Apelos por transparência desse tipo são mais comuns entre atores governamentais e ONGs.

Dessa forma A transparência é apresentada nos documentos analisados pelo Council of Europe de duas formas principais:



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

- Transparência dos algoritmos e dos dados processados, em que muitas vezes envolve a promoção da explicabilidade dos sistemas de inteligência artificial de modo que os outputs e as decisões dos sistemas são compreensíveis aos humanos.
- Transparência em relação às práticas de design e desenvolvimentos dos sistemas de modo a evitar sigilo nas estratégias de implementação da IA e conflito de interesses entre atores de IA e órgãos de supervisionamento.

UNESCO

A recomendação sobre ética da inteligência artificial apresentado pela UNESCO, em 2021, de modo similar aos documentos mencionados anteriormente destaca entre os princípios éticos a transparência e explicabilidade, indicando a sua ligação:

37. Com frequência, a transparência e a explicabilidade dos sistemas de IA são pré-requisitos essenciais para garantir o respeito, a proteção e a promoção dos direitos humanos, das liberdades fundamentais e dos princípios éticos. A transparência é necessária para que os regimes de responsabilidade nacionais e internacionais pertinentes funcionem de forma efetiva. A falta de transparência também pode prejudicar a possibilidade de se contestar efetivamente as decisões com base nos resultados produzidos pelos sistemas de IA e, com isso, infringir o direito a um julgamento justo e a um recurso eficaz, além de limitar as áreas em que esses sistemas podem ser utilizados de maneira legal.

38. Embora sejam necessários esforços para aumentar a transparência e a explicabilidade dos sistemas de IA, incluindo aqueles com impacto extraterritorial, ao longo de seu ciclo de vida para apoiar a governança democrática, o nível desses dois aspectos sempre deve ser adequado ao contexto e ao impacto, pois pode haver uma necessidade de equilíbrio entre transparência, explicabilidade e outros princípios, como privacidade, segurança e proteção. As pessoas devem ser plenamente informadas quando uma decisão é fundamentada ou tomada com base em algoritmos de IA, inclusive quando ela afeta sua segurança ou seus direitos humanos; nessas circunstâncias, elas também devem ter a oportunidade de solicitar explicações do ator de IA ou das instituições do setor público pertinentes.



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

Ademais, os indivíduos devem ter acesso às razões de decisões que afetam seus direitos e liberdades, bem como ter a opção de apresentar o caso a um membro designado da empresa do setor privado ou instituição do setor público que seja capaz de analisar e corrigir a decisão. Os atores de IA devem informar aos usuários quando um produto ou serviço for fornecido diretamente ou com o auxílio de sistemas de IA de uma maneira adequada e oportuna.

39. Do ponto de vista sociotécnico, uma maior transparência contribui para sociedades mais pacíficas, justas, democráticas e inclusivas. Ela possibilita o escrutínio público, que pode reduzir a corrupção e a discriminação, e também pode ajudar a detectar e prevenir impactos negativos sobre os direitos humanos. A transparência tem como objetivo o fornecimento de informações adequadas aos respectivos destinatários, informações essas que permitam sua compreensão e fomentem a confiança. No caso específico dos sistemas de IA, a transparência pode permitir que as pessoas entendam como cada estágio de um sistema é colocado em prática, conforme o contexto e a sensibilidade do sistema de IA. Ela também pode incluir informações sobre os fatores que afetam uma previsão ou decisão específica, e se garantias apropriadas (como medidas de segurança ou de justiça) estão ou não em vigor. Em casos de ameaças graves de impactos adversos aos direitos humanos, a transparência também pode exigir o compartilhamento de códigos ou conjuntos de dados.

40. A explicabilidade significa tornar inteligível e fornecer informações sobre o resultado dos sistemas de IA. A explicabilidade desses sistemas também se refere à compreensibilidade sobre a entrada, a saída e o funcionamento de cada bloco de construção dos algoritmos e como ele contribui para o resultado dos sistemas. Assim, a explicabilidade está estreitamente relacionada com a transparência, pois os resultados e os subprocessos que conduzem a resultados devem almejar serem compreensíveis e rastreáveis, conforme o contexto. Os atores de IA devem se comprometer a garantir que os algoritmos desenvolvidos sejam explicáveis. No caso de aplicativos de IA que afetam o usuário final de uma forma que não seja temporária, facilmente reversível ou de baixo risco, deve-se garantir que seja fornecida uma explicação significativa para qualquer decisão que resultou na ação tomada, para que o resultado seja considerado transparente.

41. Transparência e explicabilidade estão intimamente relacionadas com medidas de responsabilidade e prestação de contas adequadas, assim como com a confiabilidade dos sistemas de IA.

Do exposto há que se colocar as diferenças de percepções entre as três últimas instituições listadas:



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

○ **OCDE** – Transparência e explicabilidade são apresentados em conjunto, sem diferenciação clara. E a explicação sobre a questão engloba o acesso a informação dos sistemas de IA para o entendimento geral sobre os sistemas, para conscientização da interação com sistemas de inteligência artificial e para a compreensão do resultado e contestação.

○ **GPAN** - A transparência é apresentada como princípio. E a explicabilidade é apresentada relacionada a transparência, como forma de alcançá-la. Assim, tal princípio engloba informações sobre os dados, os sistemas e os modelos de negócios, sendo indicadas como importantes para o alcance da transparência: a rastreabilidade, como forma de auxiliar na identificação do motivo pelo qual os sistemas foram equivocados, prevenindo futuros erros; a explicabilidade, sendo o entendimento pelos indivíduos sobre como as decisões foram feitas pelo sistema de IA, devendo tal explicação ser adaptada ao nível de entendimento do interessado.

○ **UNESCO** - a transparência e a explicabilidade são apresentadas de forma mais distinta. A transparência implica o entendimento dos estágios do sistema de IA, podendo incluir entendimento sobre os fatores que afetam um previsão ou decisão, bem como, em caso de séria ameaça de impactos negativos, pode englobar o compartilhamento de códigos e dados. Já a explicabilidade é apresentada como a compreensão sobre os inputs, outputs e funcionamento dos sistemas, sendo destacado a relação de tal princípio à transparência e de que resultados devem ser compreensíveis, rastreáveis e adequados.

Ademais, como indicado pelo estudo do Council of Europe, as diretrizes apresentadas sobre inteligência artificial em sua maioria não indicam com clareza a forma pela qual os princípios da transparência e da explicabilidade serão alcançados. Além disso, a escolha dos princípios da transparência e/ou explicabilidade ainda são pauta de certa divergência nos artigos científicos sobre a questão.



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

Pasquale (2015, p. 8)¹² entende que a transparência, sendo entendida como o acesso a informação técnica de como foi feito o processo decisório, por exemplo, com o compartilhamento dos códigos, podem ser utilizados para manter a opacidade dos processos decisórios, tendo em vista que a informação não seria entendida por grande parte das pessoas devido a tecnicidade e complexidade da informação.

[...] Their methods are hard to understand primarily because of a mix of real and legal secrecy, and their scale. Interlocking technical and legal prohibitions prevent anyone outside such a company from understanding fundamental facts about it. Activists often press for transparency as a solution to the black box issues raised in this book. In many cases, sunshine truly is the “best disinfectant.” **However, transparency may simply provoke complexity that is as effective at defeating understanding as real or legal secrecy.** Government has frequently stepped in to require disclosure and “plain language” formats for consumers. **But financiers have parried transparency rules with more complex transactions. When this happens, without substantial gains in efficiency, regulators should step in and limit complexity. Transparency is not just an end in itself, but an interim step on the road to intelligibility.**

¹² Pasquale, Frank. The black box society : the secret algorithms that control money and information. London: Harvard University Press, 2015.



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

3.3 Revisão e Direito à intervenção humana

Partindo do pressuposto de que um dos objetivos da Inteligência Artificial é de alguma forma reproduzir atos de inteligência nos mais diferentes contextos de aplicação, é possível compreender que um dos usos principais da IA é voltado para a automação de decisões. Isso porque, a maior parte das atividades humanas cotidianas envolvem decisões, e o corrente desenvolvimento tecnológico traz a possibilidade de simplificar e celebrar processos mediante o uso da máquina.

Nesse sentido, vale o esclarecimento de Emre Bayamlioglu (2018, p. 6), sobre decisões:

Embora possa parecer autoexplicativo, o termo "decisão" requer uma interpretação ampla de modo a abranger qualquer ato ou resultado que mude ou tenha o potencial de mudar o status social, psicológico, econômico ou legal do indivíduo - incluindo conceitos semelhantes e termos como: planos, sugestões, propostas, conselhos ou mapeamento de opções.

Nessa toada o *Article 29 Data Protection Working Party* (2017, p. 8), em suas “Orientações sobre as decisões automatizadas e a definição de perfis”, explica que: “As decisões exclusivamente automatizadas correspondem à capacidade de tomar decisões através de meios tecnológicos e sem intervenção humana”. Ou seja, são escolhas tomadas por uma máquina, mediante manipulação de dados, que tem o potencial efetivo de alterar o estado normal das coisas.

E, exatamente por poder gerar uma alteração na vida dos titulares que são submetidos a uma atividade de decisão realizada por uma inteligência artificial, que se pressupõe a necessidade de que seja dado a eles a possibilidade de que as decisões sejam revisadas ou que tenham uma interferência humana mínima de acordo com a situação.

Para essa subdivisão especificamente foram analisados as instituições e países a seguir dispostos, com ênfase nos países Sul-americanos:

- Lei Geral de Proteção de Dados Europeia e WP-29 – Article 29 Data Protection Working Party (Grupo de Trabalho do Artigo 29º sobre a Proteção de Dados), que



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

é o grupo de trabalho europeu independente que tratou de questões relacionadas à proteção de privacidade e dados pessoais até 25 de maio de 2018 (entrada em aplicação do GDPR). Para esse tópico levamos em consideração o trabalho intitulado Orientações sobre as decisões automatizadas e a definição de perfis.

(Documento disponível em:
<https://ec.europa.eu/newsroom/article29/items/612053/en>)

- Argentina – pois é o segundo país da América Latina, após o Chile, a tratar da proteção de dados pessoais por meio da Lei 25.326, em 2000 e o Decreto regulamentar 1558/2001. Ademais, de acordo com a Decisão da Comissão Europeia n° 490/2003, a Argentina teve sua legislação reconhecida pela União Europeia como adequada para o tratamento e transmissão internacional de dados com países europeus, tendo, portanto, adequado nível de proteção de dados pessoais. Vale ressaltar que na Argentina na busca de adequar-se ao que dispunha o regulamento europeu, em 2018, foi proposto um Projeto Lei (Mensaje 147/2018) que visa alterar a atual Lei Geral de Proteção de dados

(Documentos disponíveis em:
http://www.jus.gob.ar/media/3201023/personal_data_protection_act25326.pdf /
https://www.argentina.gob.ar/sites/default/files/mensaje_nde_147-2018_datos_personales.pdf)

- Uruguai – pois foi o terceiro país na América Latina a criar sua própria legislação no que diz respeito a proteção de dados, tendo como normativa principal a Lei de Proteção de Dados Pessoais e Ação de Habeas Data (Lei n° 18.331/2008). Ademais, de acordo com a Decisão de Execução da Comissão n° 484/2012, estabelecida nos termos da Diretiva 95/46 do Parlamento Europeu, o país teve reconhecido, que assegura um nível adequado de proteção dos dados pessoais no que se refere ao tratamento automatizado de dados.

(Documentos disponíveis em: <https://www.impo.com.uy/bases/leyes/18331-2008>)



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

- Colômbia – onde mais uma vez será analisado seu Ethical Framework for Artificial Intelligence in Colombia vinculado a presidência da república, publicado em 2020.

(Documentos disponíveis em: <https://www.impo.com.uy/bases/leyes/19670-2018>)

GDPR e WP29

A Lei europeia traz expressamente a proibição de submissão a decisão automatizada sem que seja dado ao titular a possibilidade de uma intervenção humana. E o WP29 traz elucidacões relevantes sobre o que se refere uma interferência humana e como ela não pode ser fabricada.

GDPR

Artigo 22º Decisões individuais automatizadas, incluindo definição de perfis

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

2.O nº 1 não se aplica se a decisão:

- a) For necessária para a celebração ou a execução de um contrato entre o titular dos dados e um responsável pelo tratamento;
- b) For autorizada pelo direito da União ou do Estado-Membro a que o responsável pelo tratamento estiver sujeito, e na qual estejam igualmente previstas medidas adequadas para salvaguardar os direitos e liberdades e os legítimos interesses do titular dos dados; ou
- c) For baseada no consentimento explícito do titular dos dados.

3.Nos casos a que se referem o nº 2, alíneas a) e c), o responsável pelo tratamento aplica medidas adequadas para salvaguardar os direitos e liberdades e legítimos interesses do titular dos dados, designadamente o direito de, pelo menos, **obter intervenção humana por parte do responsável, manifestar o seu ponto de vista e contestar a decisão**.

4.As decisões a que se refere o nº 2 não se baseiam nas categorias especiais de dados pessoais a que se refere o artigo 9º, nº 1, a não ser que o nº 2, alínea a) ou g), do mesmo artigo sejam aplicáveis e sejam aplicadas medidas adequadas para salvaguardar os direitos e liberdades e os legítimos interesses do titular



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

WP29

O responsável pelo tratamento não pode eximir-se do disposto no artigo 22.º fabricando uma intervenção humana. Por exemplo, se alguém aplicar de forma sistemática perfis gerados automaticamente a pessoas sem ter qualquer influência efetiva no resultado, tratar-se-á, ainda assim, de uma decisão tomada exclusivamente com base no tratamento automatizado.

Para que se considere haver uma intervenção humana, o responsável pelo tratamento tem de garantir que qualquer supervisão da decisão seja relevante, e não um mero gesto simbólico. Essa supervisão deve ser levada a cabo por alguém com autoridade e competência para alterar a decisão e que, no âmbito da análise, deverá tomar em consideração todos os dados pertinentes.

No âmbito da sua AIPD, compete ao responsável pelo tratamento identificar e registar o grau de intervenção humana no processo decisório e a fase em que essa intervenção ocorre.

COLÔMBIA

O documento colombiano traz uma série de princípios relevantes que devem direcionar a regulamentação e o uso da IA, dentre eles encontra-se o da disposição a seguir:

Controle humano das decisões de um sistema de inteligência artificial

Este princípio é aplicável a sistemas de inteligência artificial que possuem certa autonomia na tomada de decisão, fazendo com que o ser humano tenha total controle sobre a tomada de decisão, principalmente na fase de implementação.

Uma vez alcançado um maior grau de maturidade da tecnologia no país, haverá maior grau de autonomia na tomada de decisões, com mecanismos de intervenção humana, principalmente quando se apresentam resultados indesejados. Essa transição levará em conta o impacto social que poderá existir, principalmente no futuro do trabalho, dado o deslocamento de seres humanos de determinadas atividades.

Na ética dos dados: a coleta e o tratamento devem ser feitos de acordo com os parâmetros e critérios estabelecidos pelos seres humanos.

Na ética dos algoritmos: os algoritmos devem permitir e facilitar a tomada de decisão, mas devem inicialmente orientar a tomada de decisão e não podem atuar de forma automatizada e segundo modelos sugeridos.



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

Na ética da prática: os sistemas de inteligência artificial não devem ser usados para interagir com os cidadãos sem o controle de um ser humano. Os sistemas automatizados de resposta e conversação devem ter mecanismos para que os humanos intervenham e participar a qualquer momento. Devem ser evitadas práticas que promovam relacionamentos com esses sistemas sem verificar se o ser humano está por trás dos conteúdos ou respostas que são gerados.

Dessa forma, é perceptível que o regulamento colombiano é favorável a interferência humana nas atividades automatizadas.

ARGENTINA

Inicialmente no que se refere a lei argente pode-se perceber um efetivo impedimento de decisões automatizadas, tornando-os nulos caso não tenham nenhum tipo de interferência humana. Vale ressaltar que em todo o texto há um forte apelo no sentido de trazer ao titular informações acessíveis em relação aos tratamentos de dados aos quais irá se submeter.

Lei 25.326

ARTIGO 20. - (Desafio de avaliações pessoais).

1. As decisões judiciais ou os atos administrativos que impliquem na apreciação ou avaliação da conduta humana, **não podem basear-se unicamente no resultado do tratamento informatizado de dados pessoais que definam o perfil ou a personalidade do interessado.**
2. Os atos contrários à disposição anterior são insanavelmente nulos.

ARTIGO 15. - (Conteúdo da informação).

1. A informação deve ser prestada de forma clara, isenta de codificação e, se for caso disso, **acompanhada de uma explicação, em linguagem acessível ao conhecimento médio da população, dos termos utilizados.**
2. A informação deve ser abrangente e abranger todo o cadastro pertencente ao titular, mesmo que o pedido inclua apenas um aspecto dos dados pessoais. Em nenhum caso o relatório poderá revelar dados pertencentes a terceiros, ainda que estejam vinculados ao interessado.
3. A informação, a critério do titular, pode ser prestada por escrito, por via eletrônica, telefone, imagem, ou outro meio adequado para o efeito.



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

No que se refere ao projeto de lei o legislador já vem com uma percepção diferente, entendendo que as decisões automatizadas e o uso da IA são inevitáveis e cada vez mais presentes no cotidiano, a intenção dele foi assegurar ao titular a possibilidade de se opor ao tratamento, além da necessidade de um relatório de impacto desses processos.

Mensaje 147/2018

ARTIGO 28. - Conteúdo da informação. **A informação deve ser prestada de forma clara, isenta de codificação e, se for caso disso, acompanhada de uma explicação dos termos utilizados, em linguagem acessível ao conhecimento médio da população**, e deve contemplar: [...] h) **A existência de decisões automatizadas**, incluindo a criação dos perfis a que se refere o artigo 32.º e, pelo menos nesses casos, informações significativas sobre a lógica aplicada, sem prejuízo dos direitos intelectuais do responsável pelo tratamento dos dados.

ARTIGO 32.- Avaliações pessoais automatizadas. **O proprietário dos dados tem o direito de se opor a ser objeto de uma decisão baseada unicamente no processamento automatizado de dados**, incluindo a criação de perfis, que produza efeitos jurídicos prejudiciais ou os afete de forma significativa e negativa. O proprietário dos dados não poderá exercer este direito se a decisão:

- a) Seja necessária para a celebração ou execução de contrato entre o titular dos dados e o responsável pelo tratamento;
- b) Está autorizado por lei;
- c) Baseia-se no seu consentimento expreso.

Nos casos referidos nas alíneas a) e c), o responsável pelo tratamento deve adotar as medidas adequadas para salvaguardar os direitos do proprietário dos dados.

ARTIGO 40.- Avaliação de impacto relacionado com a proteção de dados pessoais.

Quando o responsável pelo tratamento tenciona realizar algum tipo de tratamento de dados que, pela sua natureza, âmbito, contexto ou finalidades, é susceptível de acarretar um elevado risco de afetar os direitos dos titulares dos dados abrangidos por esta Lei, deverá realizar, antes da implementação do tratamento, uma avaliação de impacto relacionada com a proteção de dados pessoais. A avaliação de impacto relacionado com a proteção de dados é obrigatória nos seguintes casos, sem prejuízo de outros estabelecidos pela autoridade de supervisão:



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

a) Avaliação sistemática e exaustiva dos aspectos pessoais das pessoas humanas com base no processamento automatizado de dados, como perfis, e com base na qual são tomadas decisões que produzem efeitos jurídicos para as pessoas humanas ou que as afetam de forma significativa;

b) Tratamento de dados sensíveis em grande escala, ou dados relacionados a registros criminais ou de contravenção

- Esse artigo é interessante, pois a proposta da lei é trazer uma segurança preventiva ao titular, evitando a sua sujeição a um tratamento que trará danos significativos. O legislador preocupou-se em trazer a necessidade da análise de impacto antes do tratamento, tornando um fator obrigatório quando se trata de processamento automatizado.

URUGUAI

No que diz respeito especificamente a revisão das decisões automatizadas a Lei nº 18.331 estabelece como um dos principais direitos do titular a possibilidade de não serem submetidas a uma decisão unicamente automatizada referente a aspectos pessoais, que poderá trazer efeitos jurídicos significativos.

Artigo 16 (Direito de contestar avaliações pessoais) - **As pessoas têm o direito de não serem submetidas a uma decisão com efeitos jurídicos que as afetem significativamente, com base no processamento automatizado de dados que visa avaliar certos aspectos de sua personalidade, como seu desempenho no trabalho, crédito, confiabilidade, conduta, entre outros.**

A parte afetada pode contestar os atos ou decisões administrativas privadas que implicam uma avaliação do seu comportamento, cuja única base é um tratamento de dados pessoais que oferece uma definição das suas características ou personalidade. Neste caso, a parte afetada terá o direito de obter informações do Responsável pela base de dados tanto sobre os critérios de avaliação quanto sobre o programa utilizado no tratamento que serviu para adoção da decisão expressa no ato.

A regra na Lei uruguaia é exatamente a não submissão do titular a decisões automatizadas, entretanto, no caso excepcional de uma vez sujeito a esse tipo de decisão, tem garantido a possibilidade de contestação e obtenção de informações do responsável



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

pela base de dados, possibilitando a interferência humana, no que se refere aos critérios adotados e o programa utilizado para a tomada de decisão.



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

3.7 Segredos Comercial e Industrial

O art. 2º, I, do Projeto de Lei 21/20 define Inteligência Artificial como “o sistema baseado em processo computacional que pode, para um determinado conjunto de objetivos definidos pelo homem, fazer previsões e recomendações ou tomar decisões que influenciam ambientes reais ou virtuais”

Principais Críticas: Regulamentação por parte da legislação brasileira em relação à propriedade industrial sobre a tecnologia.

O legislador destaca no art. 6º, IV do PL 21/20 a “transparência e explicabilidade” como um dos princípios para “o uso responsável de inteligência artificial no Brasil”

O segredo industrial tem a sua premissa baseada na construção de uma blindagem quanto ao tipo de informação da empresa, que venha trazer inovação e maior vantagem em relação à sua concorrência. Desta forma, encontram-se inseridas nessa proteção as patentes de invenção e de modelo de utilidade, o desenho industrial e o registro de marca empresarial. Para somar com esse ambiente de tranquilidade e segurança para a propriedade industrial no marco legal da IA, o inciso V do seu art. 6º que elege a “segurança” como outro princípio, ressalva a sua compatibilidade com os padrões internacionais sobre todo o processo de vida do sistema de IA.

O respeito aos princípios oriundos de tratados internacionais em que o Brasil é parte signatária é fato louvável. Como assim resta insculpido no parágrafo único do art. 6º do PL 20/21. Em nossa atual legislação em vigor, destaca-se os tratados internacionais de Haia, Nice, Cingapura e Viena. Com isso, o País reforça ainda mais a essência da proteção industrial no marco legal da inteligência artificial.

A inclusão dos tratados internacionais junto ao marco legal da inteligência artificial traz a segurança jurídica necessária para que as empresas desenvolvam suas



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

criações em solo nacional, não fugindo assim para outros países. Outrossim, resta compreender como outras legislações estão regulamentando a mesma questão.

COMISSÃO EUROPEIA

Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de Inteligência Artificial (Regulamento Inteligência Artificial) e altera determinados Atos Legislativos da União

O aumento das obrigações de transparência também não afetará desproporcionadamente o direito à proteção da propriedade intelectual (artigo 17.º, n.º 2), uma vez que estarão limitadas às informações mínimas necessárias para as pessoas singulares exercerem o seu direito à ação e à transparência necessária perante as autoridades de supervisão e execução, em conformidade com os mandatos destas. Qualquer divulgação de informações será realizada de acordo com a legislação aplicável, incluindo a Diretiva (UE) 2016/943 relativa à proteção de know-how e de informações comerciais confidenciais (segredos comerciais) contra a sua aquisição, utilização e divulgação ilegais. Quando precisam de obter acesso a informações confidenciais ou a código-fonte para analisarem o cumprimento das obrigações substanciais, as autoridades públicas e os organismos notificados ficam sujeitos a obrigações de confidencialidade vinculativas.¹³

LEGISLAÇÃO	DISPOSIÇÃO NORMATIVA
Resolução do Parlamento Europeu, de 20 de outubro de 2020, sobre os direitos de propriedade intelectual para o	1. As autoridades nacionais competentes e os organismos notificados envolvidos na aplicação do presente regulamento devem respeitar a confidencialidade das informações e dos dados obtidos no exercício das suas funções e atividades de modo que protejam, em especial:

¹³ <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:52021PC0206&from=EN>



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

desenvolvimento de tecnologias ligadas à inteligência artificial https://www.europarl.europa.eu/doceo/document/A-9-2020-0176_PT.html	a) Os direitos de propriedade intelectual e as informações comerciais confidenciais ou segredos comerciais de uma pessoa singular ou coletiva, incluindo o código-fonte, exceto nos casos a que se refere o artigo 5.º da Diretiva 2016/943 relativa à proteção de know-how e de informações comerciais confidenciais (segredos comerciais) contra a sua aquisição, utilização e divulgação ilegais;
---	--

WIPO Conversation Inteligência Artificial (IA) na política de PI

A WIPO realizou a Primeira Sessão da WIPO Conversa on IP and AI em setembro de 2019 para discutir o impacto da AI na política de IP, com o objetivo de formular coletivamente as perguntas que os formuladores de políticas precisam fazer.

Em dezembro de 2019, a OMPI publicou um rascunho do Documento de Questões sobre política de PI e IA e iniciou um processo de consulta pública solicitando feedback para ajudar a definir as questões mais prementes que os formuladores de políticas de PI enfrentarão à medida que a IA aumenta em importância. Mais de 250 propostas foram recebidas no processo de consulta.¹⁴

Em maio de 2020, foi publicado um Documento de Questões revisado sobre política de PI e IA, levando em consideração as submissões recebidas.¹⁵

Uma Segunda Sessão da Conversa da WIPO sobre IP e IA ocorreu em julho de 2020, discutindo algumas das questões estabelecidas no Documento Revisado. Mais de 2.000 pessoas de 130 países, incluindo representantes dos Estados Membros, organizações acadêmicas, científicas e privadas participaram da reunião em formato virtual.

¹⁴ https://www.wipo.int/about-ip/en/artificial_intelligence/policy_submissions.html

¹⁵ https://www.wipo.int/meetings/en/doc_details.jsp?doc_id=499504



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

A OMPI realizou a Terceira Sessão da Conversa sobre IP e IA em 4 de novembro de 2020, em formato virtual com mais de 1.500 participantes registrados de 133 países.¹⁶

RASCUNHO DO DOCUMENTO DE QUESTÕES SOBRE POLÍTICA DE PI E IA	DISPOSIÇÃO NORMATIVA
Segredos Comerciais https://www.wipo.int/about-ut-ip/en/artificial_intelligence/policy.html	<p>40. Segredos comerciais são direitos de PI que protegem informações secretas, que tenham valor pessoal ou comercial, transmitida em contexto confidencial e que sejam objeto de um esforço razoável para protegê-lo. Em geral, as diferentes abordagens jurídicas nacionais que existem com respeito aos segredos comerciais não estão harmonizadas e seus fundamentos baseiam-se na legislação sobre responsabilidade civil, privacidade, confidencialidade ou competição injusta.</p> <p>41. Embora se possa argumentar que os segredos comerciais não constituem direitos absolutos de PI no sentido tradicional, eles geralmente podem ser uma opção atraente para os titulares de direitos de PI nos casos em que a informação não pode ser protegida por direitos de PI. A falta de divulgação oferece uma vantagem comercial. Tendo em conta que os dados desempenham um papel fundamental no desenvolvimento e aplicação da IA, a rápida evolução do domínio da IA, no âmbito do qual não se pode prever o pagamento de taxas de registro de direitos de PI, e a fato de a IA ser uma tecnologia geralmente difícil de fazer engenharia reversa, os segredos comerciais estão sendo cada vez mais usados para proteger o investimento na compilação e desenvolvimento da tecnologia, manutenção de dados e inovações no campo da IA.</p> <p>42. Por um lado, a utilização de segredos comerciais no domínio da IA incentiva a inovação na IA e proporciona um enquadramento e segurança jurídica no intercâmbio de informações e dados controlados entre determinadas partes. Dada a natureza muitas vezes complexa das cadeias de valor globais e o número de entidades relacionadas à IA, como desenvolvedores de software, cientistas da computação, instituições de pesquisa, entidades privadas, produtores de dados e hosts de dados, os segredos comerciais fornecem uma base para o compartilhamento cuidadoso de informações confidenciais e para impulsionar a colaboração. Por outro lado, a falta de divulgação contribui para o chamado problema da caixa preta e pode ser um obstáculo ao livre intercâmbio de dados.</p>

¹⁶ https://www.wipo.int/meetings/en/details.jsp?meeting_id=59168



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

	<p>43. Embora a legislação sobre segredos comerciais ou proteção de informações confidenciais levante muitas questões de natureza geral, questões específicas dessa área da política de IA e PI são:</p> <p>Questão 14: Segredos Comerciais</p> <p>i) A legislação atual sobre segredos comerciais alcança o equilíbrio certo entre a proteção das inovações de IA e os interesses legítimos de terceiros em obter acesso a determinados dados e algoritmos?</p> <p>ii) A proteção de segredo comercial deve ser conferida a dados e aplicativos de IA, ou há interesse social ou ético que a proteção de segredo comercial existente não se aplique neste contexto?</p> <p>iii) Se os dados e aplicativos de IA não devem ser protegidos por segredo comercial, tal exceção deve ser limitada a certas áreas de IA, como dados e aplicativos usados na tomada de decisões judiciais?</p> <p>iv) Se os dados e aplicativos de IA não devem ser protegidos por segredo comercial, os dados e aplicativos de IA podem ser protegidos por outros direitos de PI?</p> <p>v) Se os dados e aplicações de IA fossem protegidos pelo sigilo comercial, haveria necessidade de estabelecer um mecanismo para dar valor probatório aos segredos comerciais, bem como mecanismos práticos para preservar sua confidencialidade?</p> <p>vi) Dada a importância e o escopo das aplicações de IA globalmente, há necessidade de harmonização internacional das leis de segredos comerciais?</p> <p>vii) O uso de segredos comerciais neste contexto pode levar a consequências, intencionais ou não, de parcialidade ou confiabilidade em aplicações de IA, uma vez que o uso de segredos comerciais pode aumentar a incapacidade de reprodução e explicação da IA?</p>
--	--

AMÉRICA LATINA

PAÍS	MEDIDAS LEGISLATIVAS DE PI ADAPTADAS PARA IA	ESTRATÉGIA NACIONAL OU REGIONAL DE IA COM COMPONENTE PI
ARGENTINA	Atualmente nenhuma informação disponível.	Atualmente nenhuma informação disponível.
URUGUAI	Atualmente nenhuma informação disponível.	Atualmente nenhuma informação disponível.



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

MÉXICO	Atualmente nenhuma informação disponível.	Objetivos Prioritários, Programa Setorial da Economia 2019-2024 (PROSECO) 2019 Propõe-se melhorar a eficiência dos procedimentos de direitos de propriedade industrial promovendo, entre outros, a sua acessibilidade com recurso a ferramentas tecnológicas. https://lopezobrador.org.mx/wp-content/uploads/2019/05/PLAN-NACIONAL-DE-DESARROLLO-2019-2024.pdf
CHILE	Atualmente nenhuma informação disponível.	Política Nacional de IA e Plano de Ação de IA (projeto) https://www.minciencia.gob.cl/politicaia/ O Eixo (iii) do rascunho contém uma seção sobre Propriedade Intelectual e IA.
COLÔMBIA	Atualmente nenhuma informação disponível.	Atualmente nenhuma informação disponível.

EUROPA

PAÍS	MEDIDAS LEGISLATIVAS DE PI ADAPTADAS PARA IA	ESTRATÉGIA NACIONAL OU REGIONAL DE IA COM COMPONENTE PI
ALEMANHA	Alteração da Lei de Direitos Autorais Alemã (UrhG) 2018 https://wipolex.wipo.int/en/legislation/details/17676 Em 1º de março de 2018, a Alemanha promulgou uma exceção de direitos autorais para TDM para fins científicos não comerciais na Seção 60d da UrhG. Os artigos 4 e 5 da nova Diretiva de	Estratégia de Inteligência Artificial https://www.ki-strategie-deutschland.de/home.html .



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

	Direitos Autorais DSM (UE) 790/2019 agora fornecem limitações e exceções TDM, que são em parte mais amplas do que a seção 60d UrhG em sua versão atual. A Diretiva foi transposta para a lei nacional em 7 de junho de 2021.	
FRANÇA	Atualmente nenhuma informação disponível.	Componente econômica da Estratégia Nacional de Inteligência Artificial, Inteligência Artificial ao Serviço das Empresas 2019. Esta foi desenvolvida pelo Coordenador Econômico Nacional. A propriedade intelectual é uma questão importante e um dos três eixos da estratégia. https://www.entreprises.gouv.fr/numerique/intelligence-artificielle-au-service-des-entreprises
ITÁLIA	Atualmente nenhuma informação disponível.	Atualmente nenhuma informação disponível.
REINO UNIDO	A Lei de Direitos Autorais, Projetos e Patentes de 1988 https://wipolex.wipo.int/en/legislation/details/20592 Parte 1, Capítulo 1 da Lei de Direitos Autorais, Projetos e Patentes de 1988 estabelece a subsistência, propriedade e duração da proteção de direitos autorais no Reino Unido; Os artigos 9.º, n.º 3, e 12.º, n.º 7, referem-se especificamente a obras geradas por computador. https://wipolex.wipo.int/en/legislation/details/20651 A Seção 214(2) do Copyright, Designs and Patents Act 1988 contém disposições semelhantes em relação ao direito de design não registrado do Reino Unido. Registered Designs Act	AI Sector Deal 2018 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/702810/180425_BEIS_AI_Sector_Deal_4_.pdf O relatório considera melhorar a transferência de propriedade intelectual. Recomenda que as universidades garantam que o objetivo principal de sua abordagem à comercialização da pesquisa em IA seja a exploração da propriedade intelectual (PI) e sua proteção. O IPO do Reino Unido é responsável pela IA na política de PI. Estratégia Nacional de IA https://www.gov.uk/government/publications/national-ai-strategy



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

	1949. A Seção 2(4) da Registered Designs Act 1949 estabelece que: "desenho feito será considerado o autor".	
MALTA	Estratégia Nacional de IA de Malta 2019 https://malta.ai/wp-content/uploads/2019/11/Malta_The_Ultimate_AI_Launchpad_vFinal.pdf	Em outubro de 2019, o governo maltês publicou a Estratégia Nacional de IA de Malta. O objetivo de Malta é obter uma vantagem competitiva estratégica na economia global no campo da IA..

AMÉRICA DO NORTE

PAÍS	MEDIDAS LEGISLATIVAS DE PI ADAPTADAS PARA IA	ESTRATÉGIA NACIONAL OU REGIONAL DE IA COM COMPONENTE PI
ESTADOS UNIDOS	Atualmente nenhuma informação disponível.	Atualmente nenhuma informação disponível.
CANADÁ	Atualmente nenhuma informação disponível.	Atualmente nenhuma informação disponível.

CONCLUSÕES

Tópico 1.2: Aspectos socio-técnicos

1. A aplicabilidade da IA é vasta.
2. Há tantas vantagens quanto desvantagens
3. A relação vantagem / desvantagem depende do setor ao qual a IA será aplicada.
4. As desvantagens são capazes de comprometer decisivamente a segurança social.



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

5. Portanto, a regulação é necessária, quer de modo generalizado, quer de modo setorizado.

Tópico 1.4: Princípios e objetivos

6. Segundo uma análise empírica de uma amostra que agregava órgãos governamentais, não governamentais e empresas, verifica-se a imputação, geralmente simultânea, de pelo menos seis princípios de regulação às atividades que envolvem IA:
 - a. Equidade (Fairness)
 - b. Confiabilidade e Segurança (Reliability & Safety);
 - c. Impacto Social (Social Impact);
 - d. Responsabilidade (Accountability);
 - e. Privacidade & Segurança (Privacy & Security);
 - f. Transparência (Transparency).

Tópico 3.1: Transparência e Explicabilidade

7. Estes princípios procuram assegurar o uso social da IA
8. A nossa pesquisa verificou o conteúdo desses princípios nas seguintes organizações e países:
 - a. OCDE, GEPAN, CAHAI, CAI, UNESCO
 - b. Colombia, Uruguay e Chile
9. O princípio da transparência pode ser desdobrado em dois, cada um com suas exigências específicas:
 - c. A transparência de algoritmos e métodos de processamento de dados
 - i. A exposição dos preconceitos e vieses
 - ii. A disponibilização pública dos algoritmos para análise
 - iii. A informação sobre os insumos, a operacionalidade, a finalidade e os impactos dos algoritmos, sem que isso comprometa a confidencialidade de que eles necessitam enquanto produtos



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

- d. A transparência de práticas humanas relacionadas aos algoritmos
 - i. A divulgação das pessoas envolvidas na criação e na manutenção das aplicações inteligentes
 - ii. Exposição e combate dos conflitos de interesse entre os atores da IA e os órgãos de supervisão
 - iii. Acesso às razões de decisões que afetem os direitos dos usuários
- 10. O princípio da explicabilidade exige que:
 - e. As informações técnicas ou de qualquer natureza destinadas ao público sejam claras, precisas e compreensíveis
 - f. O funcionamento das aplicações seja não só aparente, mas inteligível
 - g. Os resultados de aplicações sejam previsíveis ou rastreáveis
- 11. O grau de necessidade de transparência e explicabilidade depende em grande medida do contexto e da gravidade das consequências de um resultado errado ou inexato
- 12. Esses princípios são sempre previstos na mesma oportunidade, embora a relação entre eles possa ser muito diversa, com algumas organizações derivando a explicabilidade da transparência (GPAN) e com outros separando-os (UNESCO), além daqueles que não os diferenciam a esse ponto (OCDE).

Tópico 3.3: Revisão e Direito à Intervenção Humana

- 13. O princípio da revisão diz respeito à possibilidade não só de rever, mas de reverter as decisões tomadas pelas aplicações inteligentes
- 14. O princípio da intervenção humana diz que deve haver a presença – mínima ou máxima, a depender da situação – de um agente humano em toda aplicação inteligente. Essa presença humana deve abranger tanto os momentos de produção quanto os de manutenção das aplicações.

Tópico 3.7: Segredos Comercial e Industrial



Grupo de Estudos em Tecnologia, Informação e Sociedade
Universidade de Fortaleza (UNIFOR)

15. Estão relacionados aos direitos civis e constitucionais de proteção da propriedade intelectual.
16. Procuram restringir a transparência e a explicabilidade ao mínimo necessário para que elas se vejam cumpridas sem que afetem os direitos de propriedade e os direitos comerciais sobre as aplicações inteligentes
17. Incentivam a inovação na medida em que oferecem, pela privacidade e pela confidencialidade, vantagens comerciais.