

Cadernos do ILP

ENSINO – PESQUISA – EXTENSÃO CULTURAL
volume 4, número 5 (especial), 2023





Inteligência Artificial e Regulação Estatal

Cadernos do ILP

ENSINO – PESQUISA – EXTENSÃO CULTURAL

São Paulo, volume 4, número 5 (especial), 2023

ASSEMBLEIA LEGISLATIVA DO ESTADO DE SÃO PAULO

MESA DIRETORA

Presidente: André do Prado
1º Secretário: Teonílio Barba
2º Secretário: Rogério Nogueira
1º Vice-Presidente: Gilmaci Santos
2º Vice-Presidente: Milton Leite Filho
3º Vice-Presidente: Helinho Zanata
4º Vice-Presidente: Rafael Silva
3º Secretário: Léo Oliveira
4º Secretário: Gil Diniz

INSTITUTO DE ESTUDOS, CAPACITAÇÃO E POLÍTICAS PÚBLICAS DO PODER LEGISLATIVO DO ESTADO DE SÃO PAULO (INSTITUTO DO LEGISLATIVO PAULISTA – ILP)

Diretora-Presidente: Agnes Thais Sacilotto
Diretor Executivo: Airton Germano da Silva
Diretor Executivo: Natacha Souza Jones

CADERNOS DO ILP: ENSINO – PESQUISA – EXTENSÃO CULTURAL

Editora: Any Ortega

CORPO CONSULTIVO

Alex Peloggia, Ana Carolina Corrêa da Costa Leister, Douglas Libório,
Érika Rigotti Furtado, Gazy Andraus, Giuliano Tierno de Siqueira,
Helen Barbosa Raiz, Julio de Souza Comparini, Leonardo David Quintiliano,
Marta Maria Assumpção-Rodrigues,
Sílvio Gabriel Serrano Nunes, Sirlene Arêdes
Stanley Plácido da Rosa Silva, Thiago dos Santos Dias





© copyright dos autores
© copyright desta edição: ILP

Cadernos do ILP: Ensino – Pesquisa – Extensão Cultural

Publicação acadêmica e técnico-científica do Instituto do Legislativo Paulista
Assembleia Legislativa do Estado de São Paulo
São Paulo (SP)
Volume 4, número 5 (Especial), 2023.
ISSN: 2675-8865
Editora: Any Ortega

Produção gráfica e revisão: ILP
Capa: computador “ENIAC”

FICHA CATALOGRÁFICA

Cadernos do ILP: Ensino - Pesquisa - Extensão Cultural / Assembleia Legislativa.
Instituto de Estudos, Capacitação e Políticas Públicas do Poder Legislativo do Estado de
São Paulo. – v. 4, n. 5 Especial Curso Inteligência Artificial e Regulação Estatal – São
Paulo: Instituto do Legislativo Paulista, 2023;
23 cm (novo formato).

Semestral
ISSN: 2675-8865

1. Ciência política – Periódico. 2. Políticas públicas – Periódico. I. Instituto de Estudos,
Capacitação e Políticas Públicas do Poder Legislativo do Estado de São Paulo.

As opiniões e conceitos expressos nos textos assinados, bem como as informações
fornecidas, a originalidade dos trabalhos, a citação de fontes e o adequado uso de
imagens, são de responsabilidade dos autores, não representando concepções oficiais
desta publicação ou de sua editoria, do ILP ou da ALESP.

Publicação de livre acesso e distribuição gratuita. É permitida a reprodução, para fins
de pesquisa e educacionais, não lucrativos, desde que citada a fonte.

**Instituto de Estudos, Capacitação e Políticas Públicas do Poder Legislativo do
Estado de São Paulo (Instituto do Legislativo Paulista – ILP)**

Avenida Pedro Álvares Cabral, 201 CEP 04097-900 São Paulo (SP) – Brasil
(11) 3886-6288 / ilp@al.sp.gov.br / www.al.sp.gov.br/ilp

Editorial

O tema da Inteligência Artificial tem se mostrado como um dos mais interessantes – e inquietantes – debates da sociedade contemporânea, profundamente marcada pelo desenvolvimento das telecomunicações e da eletrônica e pelo impacto das tecnologias informacionais. A ALESP, no cumprimento de suas funções legislativas e fiscalizatórias, não poderia deixar de incentivar debate acerca do papel do Estado nessa questão.

O curso “Inteligência Artificial e Regulação Estatal”, coordenado pelo professor Leonardo Quintiliano, foi ministrado pelo ILP entre maio e junho de 2023, com carga horária de 16h. Foram abordados, no programa, os seguintes temas: noções gerais de Inteligência Artificial (conceitos, classificações, debates e aplicações); regulação jurídica e competências dos estados e municípios; transparência e explicabilidade da IA; reconhecimento facial e regulação; mineração de processos na administração pública e *compliance*; relações de trabalho e IA; ISO e Inteligência Artificial - guias e recomendações. As aulas estão disponíveis no canal do ILP no YouTube.

O objetivo central do curso foi, desse modo, introduzir conceitos sobre Inteligência Artificial em diversos campos de aplicação que podem atrair a regulação estatal, bem como discutir a competência estatal nos três níveis da Federação para regulação da IA.

O curso teve a colaboração do *Center for Artificial Intelligence*, da USP, das Comissões de Advocacia Pública e de Privacidade, Proteção de Dados e Inteligência Artificial da OAB – SP, e o apoio do Sindalesp – Sindicato dos Servidores Públicos da Assembleia Legislativa e do Tribunal de Contas do Estado de São Paulo.

Este volume especial dos *Cadernos do ILP: Ensino – Pesquisa – Extensão Cultural* traz artigos especialmente produzidos pelos professores do curso, abordando a totalidade da temática desenvolvida e possibilitando, como é uma das funções destes “Cadernos”, o aprofundamento das discussões propiciadas pelos eventos promovidos pelo Instituto. A todos os colaboradores, que se dispuseram a compartilhar conhecimento nesta iniciativa de relevante interesse público, nossos agradecimentos.

Any Ortega
Editora

Sumário

Apresentação

A REGULAÇÃO DA INTELIGÊNCIA ARTIFICIAL

Leonardo David Quiniliano

13-16

REGULANDO A INTELIGÊNCIA ARTIFICIAL:

Limites e possibilidades dos modelos jurídicos

Cristina Godoy Bernardo de Oliveira

Pedro Sberni Rodrigues

Tiago Augustini de Lima

14-31

A REGULAÇÃO DA INTELIGÊNCIA ARTIFICIAL PELOS ESTADOS, MUNICÍPIOS E DISTRITO FEDERAL

Leonardo David Quintiliano

32-50

ISO E INTELIGÊNCIA ARTIFICIAL:

guias e recomendações

João Paulo Candia Veiga

Laura Simões Camargo

51-61

IA GENERATIVA:

passos essenciais para a implantação nas organizações

Fábio Correa Xavier

62-75

OS DESAFIOS DA INTELIGÊNCIA ARTIFICIAL NO SETOR PÚBLICO

Adriana Mecelis

76-87

MINERAÇÃO DE PROCESSOS E A ADMINISTRAÇÃO PÚBLICA

Sarajane Marques Peres

Marcelo Fantinato

Renata Mendes de Araujo

Fernando Henrique da Silva Costa

88-107

**PROTEÇÃO DOS SEGREDOS INDUSTRIAL E COMERCIAL DA INTELIGÊNCIA
ARTIFICIAL E A TRANSPARÊNCIA DA ADMINISTRAÇÃO PÚBLICA**

Daniela Copetti Cravo

108-118

RECONHECIMENTO FACIAL EM ESPAÇOS PÚBLICOS:

um risco para os regimes democráticos

Rodrigo Brandão

Thiago Gomes Marcilio

119-135

***PROFILING* À BRASILEIRA:**

entre riscos reais e imediatas expectativas

Cristina Godoy Bernardo de Oliveira

Rafael Meira Silva

136-149

NEURODIREITOS:

a tutela jurídica à mente humana

Camila Pintarelli

150-165

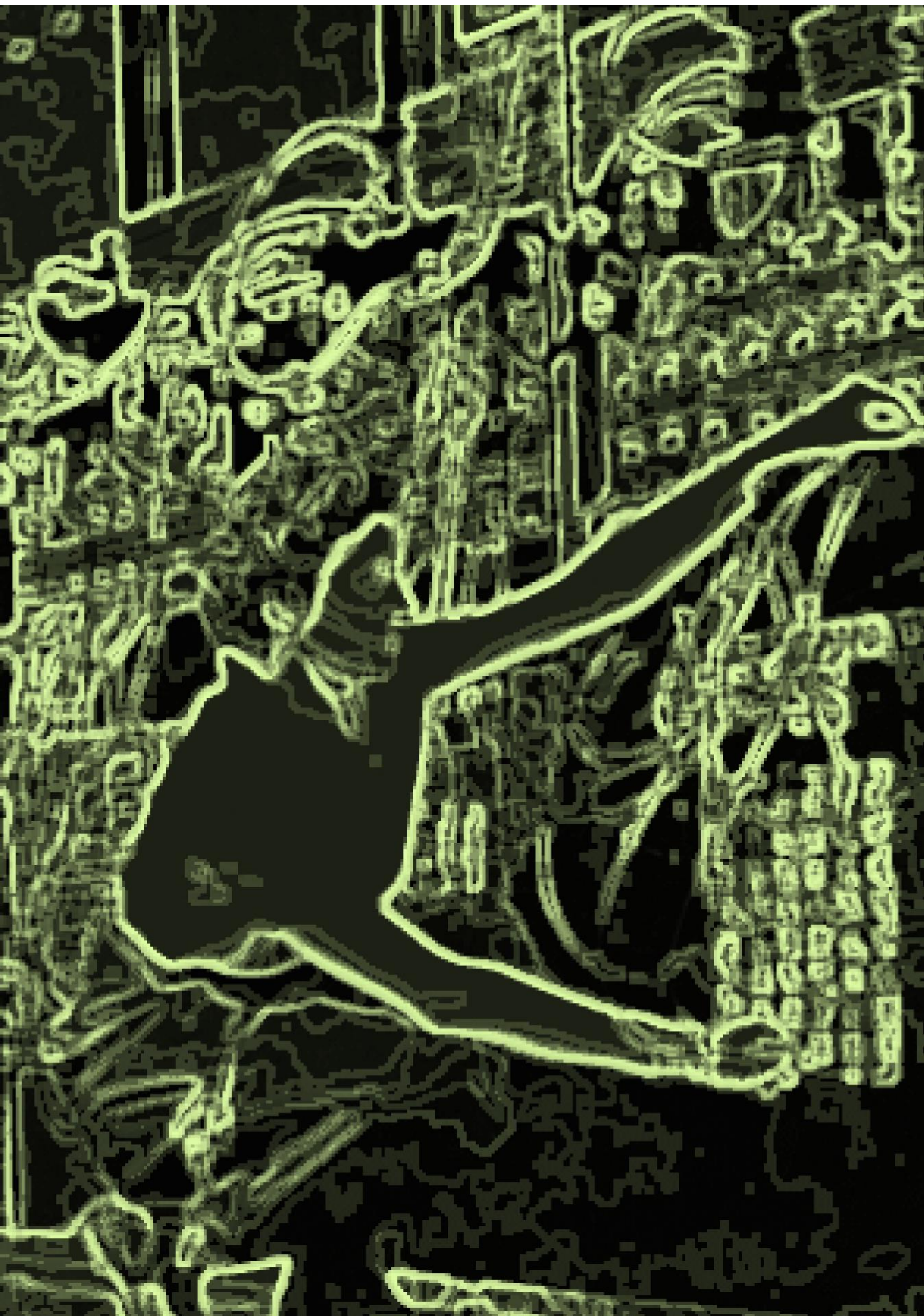
**A INTELIGÊNCIA ARTIFICIAL E O FUTURO DO TRABALHO E DA PROTEÇÃO
SOCIAL**

Alvaro A Comin

166-178

Inteligência Artificial e Regulação Estatal

Organizador: Leonardo David Quintiliano



Apresentação

A REGULAÇÃO DA INTELIGÊNCIA ARTIFICIAL

Leonardo David Quintiliano

Em 1950, o matemático inglês Alan Turing afirmou que as máquinas, um dia, poderiam competir com os homens em todos os campos intelectuais.¹ Turing estava tão convencido de sua afirmação, que chega a discutir por que campo deveriam começar os testes com inteligência artificial (xadrez, idiomas etc.), além de propor um teste para saber se as máquinas poderiam emular o pensamento humano – o jogo da imitação.²

Naquela época, contudo, a tecnologia ainda não havia experimentado um avanço que permitisse comprovar sua afirmação. Passados mais de 70 anos, a IA ainda não é capaz de emular com perfeição a inteligência humana, mas a tecnologia já se integrou no cotidiano da sociedade da informação, trazendo soluções, conforto, eficiência e agilidade, por um lado, mas riscos diversos, por outro.

A possibilidade de violação a direitos, necessidades éticas de limitação, exigências de padronização para fins comerciais, criação de ambientes com segurança jurídica para pesquisa, desenvolvimento e exploração da inovação tecnológica com recurso à IA, reclamam dos governos alguma regulação.

¹ No artigo “*Computing machinery and intelligence*” (Computadores e inteligência). Disponível em: <https://doi.org/10.1093/mind/LIX.236.433>.

² Em síntese, o jogo da imitação propõe um teste em que o computador substitui uma pessoa A que tenta confundir uma pessoa B, que tem o objetivo de saber uma característica de A (como gênero, por exemplo).

No Brasil, a regulação da utilização de sistemas de inteligência artificial também passa a ser discutida, acompanhando-se a tendência global, como apresentado no artigo “A regulação da inteligência artificial pelos Estados, Municípios e Distrito Federal”.

Ocorre que a IA não é uma matéria ou área específica de intervenção estatal, como Direito Civil, trabalho, Direito Penal, proteção do consumidor, meio ambiente, saúde, educação, segurança pública, informática etc.. Como conceito e ferramenta tecnológica, a IA já está presente em todas essas áreas, de modo que não parece ser possível e desejável uma regulação única e autônoma em um estatuto federal.

Esse foi o espírito do curso sobre Regulação da Inteligência Artificial, promovido pelo Instituto do Legislativo Paulista em parceria com a Ordem dos Advogados do Brasil, Seção São Paulo, e com o *Center for Artificial Intelligence* da USP – C4AI. A parceria permitiu a reunião de especialistas em diversas subáreas de aplicação da IA, que demonstraram em suas aulas os conceitos em torno da IA, suas áreas de aplicação, riscos, insegurança jurídica e necessidade ou não de regulação.

No artigo “Regulando a Inteligência Artificial: Limites e possibilidades dos modelos jurídicos”, são apresentados os modelos jurídicos de regulação da IA: clássico, emergente, ético e cautelar. Os pesquisadores demonstram as vantagens e desvantagens de cada um deles, concluindo pela necessidade de regulação externa a par da autorregulação dos desenvolvedores de IA.

A aplicação multissetorial da IA requer também a descentralização de regulação externa, mormente diante da necessidade de proteção de direitos fundamentais. No Brasil, tal incumbência é tarefa comum das três esferas federativas: União, Estados, Municípios e Distrito Federal. É preciso saber, contudo, quais os limites de cada um desses entes no exercício de seu poder de regulação de usos da IA em cada área passível de sua intervenção, o que constitui temática do segundo artigo.

Uma discussão que permeia o debate da regulação diz respeito a mecanismos de autorregulação, seja no âmbito *interna corporis*, seja dentro de uma área específica de atividade sujeita a regulação coletiva, embora de natureza privada, como se dá com

matérias afeitas a conselhos profissionais ou entidades setoriais. Aqui, as dificuldades estatais de adoção centralizada de um estatuto amplo para regulação da IA pode ser auxiliada pela regulação privada através do sistema ISO/IEC, explicada e exemplificada no artigo “ISO e Inteligência Artificial Guias e Recomendações”.

Os artigos seguintes fornecem exemplos de temas que podem ser objeto de regulação por um dos entes federativos, observado o limite de sua competência constitucional. A utilização pelo poder público da IA é temática abordada nos três próximos artigos, iniciando-se pelo texto “IA Generativa: passos essenciais para a implantação nas organizações”, que apresenta as vantagens da IA generativa e como ela pode auxiliar as organizações, inclusive estatais, na melhoria da eficiência e transparência, matéria que pode ser objeto de fomento estatal e regulação, com a finalidade de garantir segurança jurídica para sua adoção, como mecanismo de proteção a direitos fundamentais ou como ferramenta para sistemas de integridade.

O artigo “Os desafios da inteligência artificial no setor público” aborda a produção de resultados de eficiência na prestação de serviço e na regulação estatal necessária para orientar e controlar seu uso. Já no trabalho “Mineração de processos e Administração Pública”, demonstra-se como a mineração de processos também aparece como recurso cada vez mais adotado pela Administração Pública estrangeira e nacional para os mais diversos escopos, que incluem a verificação de conformidade em sistemas de governo eletrônico, melhoria de processos de emissão de licenças, concessão de benefícios e até mesmo a predição de eventos, como tempo de espera para atendimento.

A utilização pela Administração Pública de sistemas de IA também enfrenta o debate em torno do conflito entre o dever de transparência e a proteção do segredo industrial e comercial, como analisado no artigo “Proteção dos segredos industrial e comercial da inteligência artificial e a transparência da Administração Pública”.

Os riscos do emprego da IA em sistemas de reconhecimento facial e *profiling* e os possíveis direitos fundamentais afetados são objeto de análise e reflexões nos artigos “Reconhecimento facial em espaços públicos: um risco para os regimes democráticos” e “*Profiling*

à brasileira: entre riscos reais e imediatas expectativas”. No mesmo caminho, a cada vez mais provável emulação de habilidades mentais traz a preocupação com a salvaguarda do que podem ser chamados de “neurodireitos”, como a proteção da privacidade mental, da identidade e autonomia pessoal, do livre arbítrio e da autodeterminação, contra o preconceito algorítmico, entre outros, temática do artigo “Neurodireitos: a tutela jurídica à mente humana”.

O volume é fechado com o artigo “A inteligência artificial e o futuro do trabalho e da proteção social”, que aborda os temas do impacto da inteligência artificial sobre os empregos e profissões e da expansão das plataformas de trabalho.

Os textos que compõem esta edição especial desenvolvem, total ou parcialmente, o conteúdo de apresentações feitas no Curso “Inteligência Artificial e Regulação Estatal,” ocorridas entre os dias 11 de maio a 29 de junho de 2023.³

A riqueza do conteúdo abordado e trazido voluntária e gratuitamente pelos profissionais convidados, bem como o sucesso do evento, só foi possível graças ao empenho, zelo e dedicação ímpares das colaboradoras do Instituto do Legislativo Paulista, Any Marise Ortega e Letícia Maria de Maia Resende, ao apoio efetuado pela diretora-presidente Agnes Thais Sacilotto e pelos diretores Airton Germano da Silva e Natacha Souza Jones, às quais se soma a coordenação conjunta do evento promovida pela Professora Doutora da Faculdade de Direito da Universidade de São Paulo (RP) e Líder do Grupo de Pesquisa Direito, Ética e Inteligência Artificial da USP – CNPq, Cristina Godoy Bernardo de Oliveira, e pela Procuradora Federal e membra efetiva da Comissão de Comissão de Proteção, privacidade de Dados e Inteligência Artificial da OAB/SP, Adriana Mecelis.

Boa leitura!

Leonardo David Quintiliano
Organizador

³ O conteúdo das apresentações pode ser acessado em: <https://www.youtube.com/playlist?list=PLJT6hymwuEx-tbfasPm3LU6K7sGpKJvEq>.

REGULANDO A INTELIGÊNCIA ARTIFICIAL: limites e possibilidades dos modelos jurídicos

Cristina Godoy Bernardo de Oliveira
Pedro Sberni Rodrigues
Tiago Augustini de Lima

Introdução

Tratar da Inteligência artificial como objeto de regulação, diante de uma perspectiva jurídica alinhada é um desafio. Discute-se, pois, a delimitação do conceito de Inteligência Artificial como ponto de partida de uma observação regulatória, questionando-se, em seguida, qual é a forma mais adequada de regular a IA de acordo com as possibilidades de modelos regulatórios jurídicos existentes e as dificuldades encontradas em cada um desses modelos em contexto e perspectivas próprias.

A conceituação de Inteligência Artificial é, por si só, alvo de amplo esforço acadêmico que resulta em uma pluralidade considerável de definições existentes. No entanto, com o objetivo de se tratar dos modelos jurídicos de regulação de IA, o presente esforço parte do lugar comum de que a IA é tida, para a análise dos modelos de regulamentação, como o campo concernente à construção de sistemas que são capazes de se comportar racionalmente em uma dada situação. (RUSSELL e NORVIG 2010).

A IA é utilizada principalmente para automatizar determinadas tarefas como instrumento para otimização de decisões, funcionando como uma construção de sistemas que se sobrepõem à ação humana em eficiência, incluindo, como disciplina, diversas técnicas e abordagens, como: *machine learning* (v.g., *deep learning*

and reinforcement learning), *machine reasoning* (v.g., *knowledge representation and reasoning search*) e robótica (v.g., *control, sensors and cyber-physical systems*).

Sistemas de IA são sistemas desenhados por humanos, de acordo com um objetivo, através da construção de algoritmos, atuando na dimensão física ou digital, percebendo seu ambiente por meio da aquisição de dados, interpretando dados estruturados ou não estruturados coletados, raciocinando sobre o conhecimento ou processando as informações derivadas desses dados e decidindo as melhores ações a serem tomadas para atingir o objetivo determinado (FACELI *et al.* 2011).

Ressaltar essa dimensão é crucial para se entender a importância do desenvolvimento de um modelo regulatório de IA, uma vez que até o presente momento não é possível falar em uma inteligência artificial geral, mas em sistemas desenhados e estruturados para a execução de determinados objetivos ou tarefas, especificados, sobretudo, pelo elemento humano de criação desses sistemas, pois não há inteligência artificial que cria outros sistemas de inteligência artificial autonomamente.

Sendo assim, passa-se à exposição dos principais modelos de regulação de IA presentes no mundo jurídico, com apresentação de suas complexidades, características, dificuldades e perspectivas de acordo com os respectivos contextos. Um ponto a se observar é que, à primeira vista, os modelos parecem distantes uns dos outros, entretanto, numa leitura mais atenta, podemos identificar bases comuns entre eles (PETIT e COOMAN 2020).

Modelos Jurídicos de Regulação da IA

A regulação de IA é, sobretudo, objeto de preocupação, pois seu desenvolvimento pressupõe cuidados éticos para impedir a ofensa a direitos e garantias fundamentais no tratamento de Dados Sensíveis, ou para evitar a discriminação algorítmica (CARVALHO 2022).

Apresentam-se, como sendo os quatro grandes modelos de regulação de IA sob o ponto de vista jurídico de acordo com a prevalência do debate acadêmico os seguintes: (i) modelo clássico; (ii) modelo emergente; (iii) modelo ético e; (iv) modelo cautelar.

Modelo Clássico

O modelo clássico é reativo, descritivo, baseado na interpretação da doutrina e na legislação. Este modelo tem confecção direta da legislação já vigente. A reatividade se dá, por exemplo, quando uma nova tecnologia ocasiona dano a alguém ou consegue criar novas formas de se criar contratos etc.. Dessa forma, o direito sempre irá reagir ao fato, é a incessante busca da melhor resolução do que é dado, aceito ou proibido, através das inúmeras possibilidades restritivas que a legislação propõe (ALEXY 2011).

A principal relevância é sua aplicação em diversos campos do Direito, principalmente, seu foco principal, nas áreas de responsabilidade civil e Direito do Consumidor, no direito civil, mas também implicações em direito penal e direitos humanos (Oliveira 2023).

Para exemplificar sua aplicação, há duas problematizações que originam debates profícuos. A primeira delas se refere ao tema dos direitos da propriedade intelectual, em relação a obras de arte criadas por IA; a segunda: quais seriam os parâmetros de análise à não obviedade de uma inovação quando ela é feita por uma IA? (PETIT e COOMAN 2020).

Assim, para resolver estas questões, é necessário recorrer à legislação vigente e o judiciário tende a se manifestar, de forma reativa, por meio da interpretação e exame de finalidade das normas existentes.

Entretanto, como analisaremos em tópico específico sobre as desvantagens deste modelo, a insegurança jurídica pode gerar danos ao desenvolvimento tecnológico e à sociedade até que se consolide uma jurisprudência adequada ao tema, sobretudo na aplicação de IA na área da saúde, em que o risco é alto, por exemplo. (OLIVEIRA 2023)

Além disso, o uso de IA na área criminal, através de reconhecimento facial ou até mesmo pelo sistema bancário, para a concessão de empréstimos consignados a idosos, pode resultar em erros motivados pela falta de banco de dados adequados para aplicação de *machine learning*, ocasionando processos discriminatórios pela própria tecnologia.

Erros e imprevisibilidade das decisões são também percebidos na interpretação dos casos pelo judiciário quando da utilização de IA. A tecnologia que não possui base de dados adequada ou ideal pode gerar algoritmos racistas ou imputar atos (lícitos ou não) a outras pessoas. Portanto, a utilização das normas já vigentes e a falta de jurisprudência consolidada sobre o tema geram problemas, sendo o modelo clássico um modo pouco efetivo para regulação de IA (OLIVEIRA 2023).

Modelo Emergente

O modelo emergente é proativo, normativo, busca novas regras para fenômenos emergentes. Assim, ele é produto do Poder Legislativo após o surgimento de novos acontecimentos em determinado contexto histórico.

Como o modelo emergente cria outra área do Direito específica, sua pretensão é tentar entender as implicações da IA em questões econômicas, sociais, éticas, científicas e suas permissões e proibições (OLIVEIRA 2023).

Para isso, este modelo é baseado na identificação, análise e mitigação de riscos, independente do grau, para desenvolver nova legislação, assim, seu início ocorre com a classificação, aplicação e implicações da IA.

O grande expoente do modelo emergente é o relatório publicado no ano de 2016, pela Universidade de Stanford⁴, que aborda como a IA pode atuar em áreas como cultura, segurança, meio ambiente e suas implicações ética e de governança, além de trazer diversas definições a respeito de inteligência artificial (PETIT e COOMAN, 2020).

Outro exemplo é o relatório publicado pela OCDE (*Framework for the Classification of AI Systems*)⁵, no ano de 2022, que

⁴Disponível em : https://ai100.stanford.edu/sites/g/files/sbiybj18871/files/media/file/ai100report10032016fnl_singles.pdf Acesso em: 23 de ago, de 2023.

⁵Disponível em : <https://www.oecd-ilibrary.org/docserver/cb6d9eca-en.pdf?expires=1692910625&id=id&accname=guest&checksum=9B3B784EC0F23CA106F7FA8B834FB634.>> Acesso: 23 de agosto de 2023.

oferece diretrizes para aplicações de IA em todas as áreas, além de recomendar a criação de políticas públicas para o desenvolvimento tecnológico, educação sobre IA e gestão dos desafios impostos por IA (OLIVEIRA 2023).

As definições de IA e, por conseguinte, suas implicações em diversas áreas, são de suma importância porque, mesmo que o desenvolvimento tecnológico avance mais rápido que a legislação, diretrizes são valiosas para quem atuará na aplicação dessa tecnologia. Classificam-se os novos riscos, os setores mais ou menos sensíveis como, por exemplo, a violação da privacidade de dados pelo setor empresarial e sua distinção da violação de dados dentro da área da saúde, portanto, a taxonomia de classificação e segmentação para regular, característica do modelo emergente, é fundamental para instituição de IA segura no país (OLIVEIRA 2023).

Dessa forma, enquanto no modelo clássico utiliza-se o Direito para dirimir complicações no uso de IA, no modelo emergente o Direito direciona o que pode ser feito para não ocorrer problemas e conflitos, tendo este modelo características otimistas tanto no desenvolvimento tecnológico, quanto na parte de governança.

Assim, os grandes temas envolvendo a atuação de IA na sociedade e sua regulação pelo modelo emergente se dão em relação a nomear o que a tecnologia de IA é, com impactos, por exemplo na utilização de carros autônomos, sua responsabilidade civil e criminal; na utilização de assistentes domésticos e virtuais como Alexa, Siri, Cortana, ChatGPT, Bard etc., a utilização de drones e da robótica em geral (PETIT e COOMAN 2020).

Modelo Ético

O modelo ético tem as características da proatividade e normatividade. É um modelo teleológico quando se trata de ética deontológica e, quando da perspectiva da ética ontológica, é consequencialista. Este modelo atribui as questões éticas como as principais para a fundamentação da regulação de sistemas de IA (OLIVEIRA 2023).

A ética é objeto filosófico desde a antiguidade clássica grega e corresponde a um campo da filosofia moral desde então. Para este

estudo, dividimos a ética normativa, que analisa problemas morais específicos, em três categorias: (i) ética das virtudes; (ii) consequencialismo e (iii) deontologia (PETIT e COOMAN 2020).

A Ética das virtudes, desenvolvida pelo filósofo Aristóteles, é aquela responsável por analisar e considerar a *eudaimonia* como fator preponderante, ou seja, a busca pela felicidade como reflexo das práticas cotidianas e qualidades individuais dos sujeitos. Os exemplos clássicos propostos por Aristóteles são a sabedoria, a justiça, a inteligência e a moderação, contudo, como na filosofia aristotélica a justiça é uma virtude universal, acompanha todas as outras virtudes, sendo a justa medida o ideal a se buscar na aplicação dessas virtudes (ARISTÓTELES 2014).

O filósofo estagirita ensina que a justa medida segue a seguinte explicitação no caso da coragem, por exemplo: a virtude está mediando a imprudência e a covardia, mas sempre atuando com justiça, ou seja, para ser corajoso, o indivíduo deve ser justo (ARISTÓTELES 2014).

Já nas aplicações da ética das virtudes no contexto da IA, a principal virtude talvez seja a transparência, porque caso haja, por exemplo, ofensas à privacidade de dados, há necessária aplicação da responsabilidade e o excesso de proteção impediria o desenvolvimento da tecnologia (PETIT e COOMAN 2020).

Já na segunda proposição, o consequencialismo, foca-se no impacto, ou seja, no custo benefício, sempre buscando dar enfoque aos impactos de determinada ação (OLIVEIRA 2023).

O consequencialismo no modelo ético, portanto, possui três perspectivas de análise conceitual: (i) a primeira se refere quando se parte de uma análise a partir do agente, o consequencialismo é egoísta quanto aos custos e benefícios; (ii) entretanto, quando a análise parte da perspectiva da sociedade, excluindo o agente, o consequencialismo possui característica altruísta e (iii) quando se analisa o impacto do objeto em regulação na sociedade como um todo, o modelo é utilitarista, ou seja, busca maximizar os benefícios e a felicidade, enquanto minimiza seus riscos (PETIT e COOMAN 2020).

O modelo ético consequencialista de regulação de inteligência artificial pode ser encontrado nos objetivos das diretrizes europeias de IA: “maximizar os benefícios dos sistemas de IA, ao

mesmo tempo em que se previnem e minimizam seus riscos⁶” (HIGH LEVEL EXPERT GROUP 2019).

Já o modelo deontológico, seguindo a filosofia kantiana, não depende da consequência de uma ação, independentemente de quem a proferir. O deontológico ético propõe uma análise da ação, caracterizando-a ética ou não; boa ou não, independentemente de seu fim, sendo assim, numa ação ou desenvolvimento de IA o que se deve buscar é a imposição de um imperativo categórico universal como, por exemplo, na maioria dos documentos jurídicos da União Europeia e, inclusive, nas diretrizes europeias de IA, quando prescrevem que a IA não possui um fim em si mesma, mas busca melhorar o desenvolvimento humano, servindo à humanidade e não o contrário (OLIVEIRA 2023).

Nicolas Petit e Jerome Cooman (2020, p.14) detalham um paradoxo que o modelo ético deontológico possui de maneira didática: os autores descrevem a situação de um assistente doméstico, como a Alexa, treinado com o imperativo categórico para sempre dizer a verdade, que é questionada por um assassino onde seu dono está. Portanto, uma pequena dose de consequencialismo faria bem.

Modelo Cautelar

Por fim, temos o modelo cautelar que é proativo, normativo, baseado no risco, no cálculo do custo-benefício e no princípio da precaução. Este é o modelo utilizado pela Comissão Europeia de regulação de IA e também pelo High Level Expert Group. O modelo cautelar de regulação propõe a minimização de riscos e modos de prevenir possíveis danos aos usuários, sendo este o modelo regulatório em pauta no Senado Federal, sob a relatoria do Senador Rodrigo Pacheco (PSD/MG) (OLIVEIRA, 2023).

O modelo cautelar é interessante porque apresenta antecipadamente propostas para correção de falhas em *design* e segurança da IA, mas também modos de responsabilização adequados,

⁶ Tradução nossa de: “This entails seeking to maximise the benefits of AI systems while at the same time preventing and minimising their risks (High Level Expert Group, 2019, p. 6).”

pois o modelo é preventivo e não corretivo, com a possibilidade de intervenção humana para dirimir ou melhor orientar o sistema de IA (PETIT e COOMAN 2020).

A regulação analisada por riscos possibilita análises estatísticas do uso de IA através de acertos e erros, ajudando a preveni-los. Um exemplo disso, já utilizado no Brasil, são os boxes regulatórios no setor financeiro. *Box* regulatório é o treinamento simulado de uma ação utilizando IA antes de sua aplicação, porque evita danos e erros. Na Alemanha é utilizado um sistema piramidal com cinco níveis de risco, desde os negligenciáveis, os que não precisam de regulação, aos existenciais, que necessitam de ser eliminados (OLIVEIRA 2023).

Assim, dentro do modelo cautelar, deve-se realizar o cálculo do risco através do princípio da precaução quando, por exemplo, as redes neurais levam o treinamento de IA à aplicação de pontuações ou qualificações de pessoas, produção de discriminações, armas de destruição em massa e etc. Portanto, mesmo no modelo cautelar são utilizados, em alguma medida, pressupostos do modelo ético (OLIVEIRA 2023).

Das vantagens e desvantagens

As principais vantagens do modelo clássico de regulação de IA são duas, a saber: (i) melhor eficiência em relação à coerção, com a possibilidade de sanção; (ii) a possibilidade de previsibilidade de decisões nas várias instâncias do Brasil (OLIVEIRA *et al.* 2022).

Contudo, há também as desvantagens e as principais são: (i) alta onerosidade e pouca flexibilidade e, em se tratando de tecnologia, sistemas legais rígidos se tornam defasados em pouco tempo; (ii) o tempo demasiadamente longo para aprovação de leis não acompanha os avanços tecnológicos e (iii) a definição de IA e suas inúmeras aplicações e implicações não são suficientes para acompanhar o desenvolvimento dessas tecnologias (OLIVEIRA *et al.* 2022).

Há de se ressaltar que recair no lugar-comum de que já existem leis suficientes é subestimar que normas jurídicas são representações da realidade e consolidação de lutas sociais históricas, retomando-se os ensinamentos da clássica obra de Rudolf Von Ihering

(1872), *A Luta pelo Direito*. É imprescindível, pois, que haja envolvimento social na produção das normas para a criação do Direito.

Em se tratando de um fenômeno emergente gerador de situações inusitadas e inovadoras, não é possível responder ao seu conteúdo fático de maneira a aplicar automaticamente um ordenamento jurídico que reproduza as noções de uma outra realidade desconexa com a causalidade dos novos fenômenos. Diante disso, verifica-se a dificuldade de estabelecer um Marco Legal da Inteligência Artificial, não se conseguindo atingir um sistema jurídico autossuficiente (OLIVEIRA *et al.* 2022).

Tal situação se exemplifica em casos de possíveis descobertas, invenções e obras artísticas realizadas por uma inteligência artificial e os direitos decorrentes dessas criações intelectuais ou em casos de utilização de base de dados eivadas de vícios - não é possível tratar o tema indefinidamente com a legislação preexistente sem que se recaia em situações teratológicas. Reitera-se, assim, o papel crucial do Poder Legislativo como legitimado constitucional para tratar dessas questões emergentes na positivação de direitos, capaz de unificar e determinar as regras e nortear a atuação do poder judiciário (OLIVEIRA 2023).

O modelo emergente, por sua vez, é caracterizado pela busca de uma estruturação de um novo campo do Direito e a criação de uma taxonomia própria. Apresenta-se como grande contribuição do modelo o esforço no sentido de se criar classificações, categorias e desdobramentos, o que permite ao legislador pautar seu trabalho sob diferentes perspectivas, como a análise de risco, impacto nos direitos humanos e afins, de maneira pormenorizada, contrapondo-se à codificação em bloco, como um marco regulatório, o que apresenta uma dificuldade muito maior de satisfazer a regulação de diversas aplicações da IA em sociedade (OLIVEIRA 2023).

A vantagem do modelo é a possibilidade de uma taxonomia, isto é, classificação e segmentação para regular de forma estrutural e satisfatória a IA em seus mais diferentes setores, não se tratando de um modelo reativo, uma vez que se busca regular questões jurídicas de maneira prévia e antecipada (OLIVEIRA 2023).

No entanto, o modelo de regulação emergente apresenta como importante dificuldade a incapacidade de capacitação nos temas

de IA por parte de profissionais de outros campos e legisladores, tendo em vista o alto grau de especificidade e a dinâmica inventiva marcada por enorme fluidez, o que gera dificuldades e ruídos na comunicação entre especialistas e leigos e dificulta a participação da sociedade civil no desenvolvimento de discussões públicas e mesmo na consolidação dos direitos (OLIVEIRA 2023).

Para além disso, a impressão de que o ordenamento jurídico positivo é insuficiente em absoluto para tratar dos novos temas de IA e que um novo ramo do Direito traria, a partir de sua taxonomia, todo um novo sistema. Não é verdade, ressalta-se que o princípio da unidade da Constituição, por exemplo, demonstra que a organicidade do ordenamento jurídico e de sua aplicabilidade são lentes importantes no desenvolvimento do Direito em suas mais variadas acepções e ramos (OLIVEIRA 2023).

Tratando-se do modelo ético, é preciso retomar a divisão da ética normativa nas três subcategorias alhures explicitadas, quais sejam, (i) ética das virtudes - trazendo a transparência e a explicabilidade como pilares de sua aplicação na IA; (ii) consequencialismo - objetivando maximizar os ganhos e diminuir os riscos dos sistemas de IA da maneira mais eficiente possível, presente nos *European Guidelines on AI* (p.4) e; (iii) deontologia, também adotada pelo *European Guidelines on AI*, no sentido de encarar a IA como um meio, um instrumento para a ação humana, não como um fim em si mesma (OLIVEIRA 2023).

Suas vantagens são claras, já que permite uma interpretação viável para regular diferentes áreas de aplicabilidade da IA, sobretudo em vistas a casos concretos. Nesse sentido, ressalta-se o prevalecimento do modelo regulatório ético de cunho deontológico na União Europeia. No entanto, diversas dificuldades e desvantagens podem ser demarcadas.

Destarte, a fixação de primados éticos dentro da deontologia pode gerar paradoxos, como o clássico primado de dizer sempre a verdade pois essa é a ação ética e gerar um dano não desejado pelo sistema, incorrendo, por exemplo, em possíveis falhas de segurança consideráveis (OLIVEIRA 2023).

Citam-se, ainda, Floridi e Cowls (2019), que apresentam a problemática do procedimento de importação de princípios de outras

áreas para a regulação jurídica da IA. O problema é exemplificado através do campo da bioética, no qual, para a regulação jurídica, exporta princípios como o da beneficência, não maleficência, autonomia e afins. No entanto, é questionável que assim o faça, uma vez que as relações entre programadores, desenvolvedores de sistemas de IA e os usuários e consumidores estabelecem uma relação diversa daquela que trata o campo da bioética, por exemplo entre o médico e seu paciente (OLIVEIRA 2023).

Questiona-se, também, o fato de que pesquisas apontam que pressupostos éticos têm pouco ou nenhum impacto na aplicação rotineira à função dos desenvolvedores de IA (FLORIDI e COWLS 2019).

Por fim, o modelo ético parte da premissa da existência de que os princípios éticos por ele assumidos são universalmente estabelecidos. A partir disso, problematiza-se a diversidade de padrões éticos existentes entre países e culturas diversas, o que pode resultar no *dumping* ético, prática de exportação de sistemas de IA com reduzida preocupação ética. Daí a União Europeia estar privilegiando mecanismos de certificação e garantias para importação de sistemas de IA advindos de outros países, especificamente aqueles reconhecidos por sua comunidade como países com baixos padrões éticos (OLIVEIRA 2023).

O modelo cautelar, pautado na regulação do risco, isto é, na produção de mecanismos de identificação de probabilidade de ocorrência de danos inerentes a uma tecnologia e sua consequente regulação de acordo com tal probabilidade é adotado pela comissão europeia, concentrando-se em minimizar riscos de dano potenciais, tal como prevê, inclusive, o projeto de regulação de IA pátrio, projeto do senador Rodrigo Pacheco (OLIVEIRA 2023)

Apresentam-se como limites desse modelo duas situações, quais sejam, (i) quando é impossível calcular qualquer previsão de probabilidade do risco por falta de certeza científica, a precaução deve prevalecer. De acordo com o princípio da precaução, trata-se de um dever moral assegurar que todo o possível seja feito para se evitar um risco de catástrofe, por exemplo, um risco inaceitável; (ii) quando fatores culturais políticos e psicológicos influenciam profundamente modelos de regulação entre diferentes partes da comunidade internacional, citando-se o complexo de “Frankenstein” da cultura

ocidental em temer o superdesenvolvimento de sistemas de inteligência artificial que se voltariam contra a humanidade como exemplo dessa pegada cultural que influencia a percepção dos indivíduos no que tange à identificação de riscos partindo de uma tecnologia nova ou um sistema de IA. (PETIT e COOMAN 2020).

Assim, diz-se que o princípio da precaução pode gerar situações de cautela extremada, o que seria uma desvantagem em que a sociedade desperdiçaria tempo e recursos ao buscar regular situações raras e de alto impacto, os chamados “cisnes negros” (TALEB 2007). O dispêndio em relação às exceções superaria às generalidades. A isso se exemplifica a responsabilização da energia nuclear de um modo geral pelo desastre nuclear de Fukushima em 2011, em todo o mundo, com a introdução de políticas de eliminação de produção de energia nuclear - ainda que a causa fosse a especificidade japonesa de estar geograficamente sujeita a abalos sísmicos.

Conclusão e perspectivas

A utilização de uma inteligência artificial ética é uma demanda cada vez maior na sociedade. O reflexo dessa demanda são projetos de lei no Senado e Câmara Federal, do Congresso Brasileiro. Entretanto, há inúmeras dificuldades e pressupostos a se considerar para regular a matéria - o primeiro aspecto - objeto do presente artigo - é a escolha do modelo regulatório.

Antes da análise e escolha do melhor (ou melhores) modelo jurídico para regular o uso de IA, é fundamental a descrição do que seja a IA e como ela se constitui - porque a tomada de decisão que uma IA realiza só é possível mediante seu aprendizado de máquina, ou seja, a qualidade, quantidade e tipos de dados que foram oferecidos ao sistema para que ele construa sua rede neural. Assim, o conceito (princípio) alocado a essas atividades de criação de IA é a transparência, dessa forma é interessante algumas explicações mais aprofundadas sobre transparência.

David Watson (2021, p.15) acentua que, para evitar opacidade e o que os cientistas denominam como “caixa preta”, a construção de algoritmos no *machine learning* deve respeitar os preceitos de (i) *fairness* - os sistemas de IA não podem ser

discriminatórios ou favorecer determinados grupos de pessoas em detrimento a outros em relação à raça, classe, gênero, idade; (ii) *accountability* - entender que as ações da empresa de tecnologia possuem responsabilidade em relação a seus desenvolvedores e (iii) transparência - em relação à IA e *machine learning*, significa tornar claros os processos de tomada de decisão do sistema de IA, permitindo que reguladores, especialistas examinem as atividades de IA, além disso, é a transparência que permite a identificação de algoritmos preconceituosos e erros intencionais ou não.

Tendo isto em vista, é fácil entender porque a regulação externa, ou seja, não apenas a autorregulação dos desenvolvedores, é essencial e porque a governança é imprescindível a essas empresas de tecnologia para que, independentemente dos modelos de regulação, possam ser realmente efetivas.

Então, a partir do exposto, o que seria regular IA? Regular é organizar um complexo sistema de termos, unidades, categorias para criar uma estrutura (complexa) para que a efetividade do objeto regulado seja mantida ou otimizada, sem impedir avanços tecnológicos e econômicos. A licitude do objeto é o que garante a não responsabilização - ou seja - regular é criar um sistemas de regras e normas que facilitem a interação do objeto com a sociedade, permitindo a resolução de conflitos da melhor maneira possível, prevenindo os danos e otimizando a eficiência (MENENGOLA *et al.* 2022).

Ao invés de se pregar a adoção irrestrita de um dos modelos de regulação apresentados, o presente exercício visa apresentar as possibilidades de regulação da IA e, concomitantemente, discutir suas minúcias.

Em qualquer dos modelos alhures discutidos, há de se ter sob foco também a proteção do processo de inovação, estipulando-se a compreensão das falácias, pontos negativos e desvantagens que permeiam os modelos para que seja possível buscar um desenvolvimento regulatório que evite tais desvios (OLIVEIRA 2023).

Não obstante, entende-se que desenvolver uma classificação dos sistemas de IA sob a perspectiva das categorias regulatórias aqui dispostas é um caminho para o tratamento da IA em cada um de seus contextos, uma vez que a amplitude da utilização da inteligência

artificial não prescinde de um olhar atento e especializado do regulador ao setor especializado de sua aplicação, em respeito às minúcias que influenciam e dizem respeito aos mais diversos confins a serem alcançados pela Revolução que a IA anuncia.

Referências Bibliográficas

- ALEXY, R. (2001) **Teoria da Argumentação Jurídica**. São Paulo: Landy,
- ARISTÓTELES (2014) **Ética a Nicômaco**. São Paulo: Edipro.
- CARVALHO, A.C.P.L.F. (2022) Ética e Regulação na Inteligência Artificial. In: *Ética e Regulação na Inteligência Artificial. Revista da Sociedade Brasileira de Computação*, Porto Alegre, vol. 47, junho, p. 7-10, 2022. Disponível em: <https://www.sbc.org.br/images/flippingbook/computacaobrasil/computa_47/pdf/CompBrasil_47.pdf>. Acesso em: 23 de ago. de 2023.
- FACELI, K.; LORENA, A.C.; GAMA, J.; CARVALHO, A.C.P.L.F. (2011) **Inteligência Artificial: Uma Abordagem de Inteligência de Máquina**. Rio de Janeiro: LTC.
- FLORIDI, L.; COWLS, J. (2019) **A Unified Framework of Five Principles for AI in Society**. *Harvard Data Science Review*: “Principles alone cannot guarantee ethical AI”.
- HIGH LEVEL EXPERT GROUP (2019). **Ethics Guidelines for Trustworthy AI**. European Commission, Disponível em: <<https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>>. Acesso em: 24 de ago. de 2023.
- IHERING, R. V. (2010) **A luta pelo Direito**. 6. ed. São Paulo: Editora Revista dos Tribunais.
- MENENGOLA, E.; GABARDO, E; GONZÁLEZ SANMIGUEL, N.N. (2023). **A Proposta Europeia de Regulação da Inteligência Artificial**. *Seqüência estudos Jurídicos Políticos*, 43(91), 1–27. <https://doi.org/10.5007/2177-7055.2022.e91435>. Disponível em: <<https://periodicos.ufsc.br/index.php/sequencia/article/view/91435>>. Acesso em: 26 de ago. de 2023.
- RUSSELL, S.J.; NORVIG, P. (2010). **Artificial Intelligence: a modern approach**. Upper Saddle River: Prentice Hall.
- OLIVEIRA, C.G.B. (2023) Regulação de IA e Modelos Jurídicos. 26 de maio de 2023. In: Reconhecimento facial: A tecnologia e os limites para sua utilização. ALESP. YouTube. Disponível em: <[Reconhecimento facial: A tecnologia e os limites para sua utilização](#)>. Acesso em: 23 de agosto de 2023.
- OLIVEIRA, C.G.B.; VEIGA, J.P.C.; COZMAN, F.G. (2022) Regulação da Inteligência Artificial: qual modelo adotar? In: *Ética e Regulação na Inteligência Artificial. Revista da Sociedade Brasileira de Computação*, v. 47,

junho, p.28-31. Disponível em: <https://www.sbc.org.br/images/flippingbook/computacaobrasil/computa_47/pdf/CompBrasil_47.pdf>. Acesso em: 23 de ago. de 2023.

PETIT, N.; COOMAN, J. (2020) Models of Law and Regulation for AI. *EUI Working Papers*. European University Institute. RSCAS 2020/63. Disponível em: <<https://cadmus.eui.eu/handle/1814/68536>>. Acesso em: 24 de ago. de 2023.

TALEB, N. (2007). **The Black Swan**: the Impact of the highly improbable. random house, 2007.

WATSON, D. (2021) **Explaining Black Box Algorithms**: epistemological challenges and machine learning solutions. Tese (Ph.D in information, communication and the social sciences). Universidade de Oxford, Oxford, . Disponível em: <<https://ora.ox.ac.uk/objects/uuid:ba743054-3eaf-41fc-98e8-841255ee24ad/files/dr55z7824>>. Acesso em: 26 de ago. de 2023.

Autores

Cristina Godoy Bernardo de Oliveira

Professora da FDRP - USP. Pós-doutora pela Université Paris I Panthéon-Sorbonne. *Academic Visitor* da University of Oxford. Doutora e graduada pela Faculdade de Direito da USP. Líder do Grupo de Pesquisa Direito, Ética e Inteligência Artificial da USP - CNPq.

Pedro Sberni Rodrigues

Advogado. Mestrando em Direito pela FDRP-USP. Bolsista de Mérito Acadêmico - Universität Passau (2022). Bolsista de Extensão Aprender na Comunidade: Informar para Democracia (USP). Membro do grupo de pesquisa em inteligência artificial C4AI (USP).

Tiago Augustini de Lima

Advogado. Mestrando em Direito pela Faculdade de Direito da USP – Campus Ribeirão Preto. Bolsista CAPES. Bolsista PET. Bolsista de Iniciação Científica PUB/USP. Bolsista de Pesquisa e Extensão Aprender na Comunidade: Informar para Democracia (USP), membro do grupo de pesquisa em inteligência artificial C4AI (USP).

A REGULAÇÃO DA INTELIGÊNCIA ARTIFICIAL PELOS ESTADOS, MUNICÍPIOS E DISTRITO FEDERAL

Leonardo David Quintiliano

Introdução

O avanço no uso da Inteligência Artificial - IA tem acendido cada vez mais o debate sobre a necessidade de sua regulação própria.

Se, por um lado, é possível recorrer a princípios e regras contidas em microssistemas jurídicos, como a disciplina de proteção ao consumidor, os Códigos Civil, Penal e legislação trabalhista, além de normas constitucionais autoaplicáveis, por outro, a indeterminação de algumas dessas normas e a falta de regramento específico podem criar insegurança jurídica para o uso da nova tecnologia.

Nesse sentido, parece haver um consenso internacional em torno da necessidade de uma regulação mínima para o uso de IA.

No âmbito da União Europeia, *v.g.*, o Parlamento Europeu aprovou, em 16 de fevereiro de 2017, uma resolução com recomendações à Comissão de disposições de Direito Civil sobre Robótica. (UNIAO EUROPEIA 2022). Mais recentemente, em 14 de junho de 2023, o mesmo parlamento aprovou a proposta de um Regulamento de Inteligência Artificial. (UNIAO EUROPEIA 2023).

Nos Estados Unidos da América, foi apresentado um projeto de lei (The National AI Initiative Act), em 2020, convertido em lei em 2021, a qual, entre outras medidas, cria o Comitê Consultivo Nacional de Inteligência Artificial (National Artificial Intelligence Advisory Committee – NAIAC), com a função de fornecer recomendações ao Presidente, ao então Gabinete de Iniciativa Nacional de IA e ao Congresso sobre tópicos como o estado atual da competitividade da IA

dos EUA, progresso na implementação das políticas de IA, entre outras. (ESTADOS UNIDOS DA AMÉRICA 2023).

Em 2020, um memorando conjunto encaminhado a todas as agências regulatórias dos Estados Unidos, em cumprimento da Ordem Executiva nº 13.859, de 11 de fevereiro de 2019, do Presidente dos EUA (ESTADOS UNIDOS DA AMÉRICA 2019), determinou que as agências federais devem evitar ações regulatórias ou não regulatórias que prejudiquem a inovação e o crescimento da IA.

Ou seja, ali nos EUA foi adotada uma concepção mais liberalizante, notadamente com foco na concorrência internacional diante da IA, especialmente considerando-se o uso da IA não apenas no comércio e produção, mas igualmente na indústria bélica.

Na China, a Cyberspace Administration of China (CAC), órgão regulador da internet, publicou neste ano um ato adotando medidas de gerenciamento de serviços de inteligência artificial generativa (Measures for the Management of Generative Artificial Intelligence Services). Destaque-se que o modelo chinês se apresenta fortemente intervencionista, responsabilizando inclusive os provedores de aplicações pelo conteúdo gerado pela inteligência artificial generativa. (HUANG 2023).

No Brasil, o Senado Federal criou uma comissão de juristas para apresentar uma proposta de regulação. O relatório final foi apresentado em 1º de dezembro de 2022, adotando uma posição equilibrada, conciliando uma “abordagem baseada em riscos com uma modelagem regulatória baseada em direitos”. (BRASIL 2022, p. 10)

A proposta teve por finalidade substituir projetos de lei em tramitação no Congresso Nacional (Projetos de Lei 5.051, de 2019, 21, de 2020, e 872, de 2021), que têm como objetivo estabelecer princípios, regras, diretrizes e fundamentos para regular o desenvolvimento e a aplicação da inteligência artificial no Brasil.

O PL 21, de 2020, trazia em sua redação original uma proposta de responsabilização mais ampla dos agentes de inteligência artificial. A proposta oferecida, que hoje também se vê contemplada no Projeto de Lei n. 2.338, de 2023, em tramitação no Senado Federal, propõe a mesma divisão de responsabilidade entre os diversos agentes de inteligência artificial, conforme o grau de risco da

ferramenta, aferida a partir de critérios também propostos no projeto, com a avaliação de um órgão público competente.

A despeito de todo esse arcabouço pré-normativo e das iniciativas estrangeiras, nota-se uma dificuldade na regulação do uso de inteligência artificial devida à natureza intrínseca, multidisciplinar e transversal da tecnologia.

Em um sistema interconstitucional ou transconstitucional, com relativização dos conceitos clássicos de soberania e territorialidade e a globalização de empresas como as *Big Techs*, é claro que se torna inexorável reconhecer que alguns princípios de regulação devem observar a tendência global, sob pena de ocasionar algum tipo de isolamento comercial ou econômico do país ou ente federativo.

Não obstante, as ferramentas de inteligência artificial espalham-se por todas as áreas da vida cotidiana e social, de modo cada vez mais acentuado e de forma até exponencial. Não é possível, assim, imaginar que um único estatuto federal ou até mesmo um tratado internacional discipline de forma exaustiva todas as consequências jurídicas que decorrem do uso da IA.

Além disso, cada povo e sociedade política conserva seu direito à autodeterminação.

No caso de Estados federativos, a transversalidade da IA abre espaços regulatórios compartilhados entre as diversas esferas de poder. No Brasil, tais esferas compreendem, além da União, o poder central, os Estados, o Distrito Federal e os Municípios, os quais podem adotar leis que afetem o uso de IA no âmbito de seu interesse e que se relacionem aos temas de sua competência.

As possibilidades de regulação pelos entes federativos subnacionais é tratada neste capítulo.

O âmbito de regulação da Inteligência Artificial e o federalismo brasileiro

O modelo de repartição de competências no federalismo brasileiro observa o critério taxativo-central, a exemplo dos Estados Unidos da América, com a fixação de competências expressas para o

poder central (União), reservando aos Estados as competências não enunciadas.

Constituindo o Brasil um federalismo de três níveis (federalismo trino ou sobreposto),⁷ contudo, os Municípios também possuem autonomia e competências próprias, enunciadas ou implícitas, especialmente sendo-lhes reservadas as competências para tratar de assuntos de interesse exclusivamente municipal ou local.

Há que se destacar, ainda, a adoção do regime federativo cooperativo, inspirado no modelo alemão, com competências comuns e concorrentes entre todos os entes federativos, de maneira cooperada, cabendo à União a adoção de normas de caráter geral e aos demais entes federativos o poder suplementar ou supletivo.

Em razão do sistema de repartição constitucional de competências adotado, é necessário saber que matéria se encontra afetada pelo uso da tecnologia de IA, a fim de se identificar a possibilidade de adoção de regulação própria no âmbito dos entes federativos subnacionais.

Conquanto se reconheça que a tecnologia de IA torna-se cada vez mais onipresente, podem-se destacar algumas áreas em que seu uso se mostra mais destacado, com potencial de atração da competência dos Estados e Municípios.

A competência privativa da União para legislar sobre informática

Segundo o artigo 22, IV, da Constituição Federal, compete privativamente à União legislar sobre informática. A expressão informática, contudo, abrangeria integralmente a inteligência artificial?

A Informática pode ser conceituada como a matéria que diz respeito à aplicação da informação automatizada. Sobre a origem do termo e o alcance semântico atual, parece existir alguma divergência na literatura.⁸

⁷ Cf. nossa proposta de tipologia federativa contemporânea. (QUINTILIANO 2012, p. 64-66).

⁸ Há autores que enfatizam a origem francesa da expressão, com a palavra *informatique*, derivada da junção dos vocábulos *information* e *automatique*. Embora

Apesar dessas parcialmente divergentes explicações para a origem do termo, seu significado atual parece convergir para o conceito acima, designando o estudo e aplicação da informação automatizada, como descreve o Professor de Documentação da Faculdade de Biblioteconomia e Informação Científica da Universidade de Brasília, em artigo publicado em 1968 (VICENTINI 1968).⁹

O campo temático da *informática* abrange, assim, toda informação processada automaticamente. Segundo o artigo 4º, I, da Lei de Acesso à Informação (Lei Federal nº 12.527/2011), *informação* pode ser conceituada como “dados processados ou não, que podem ser utilizados para produção e transmissão de conhecimento, contidos em qualquer meio, suporte ou formato”.

Resta saber se o âmbito de regulação da inteligência artificial coincide com a matéria *informática* ou está totalmente nela inserido. Se afirmativa a resposta, a conclusão é de que não cabe aos Estados e Municípios qualquer espaço legislativo para regulação do uso da inteligência artificial, uma vez que o tema seria objeto de tratamento privativo pela União, a teor do artigo 22, IV, da CF.

Como já abordado nesta obra, a inteligência artificial é um conceito que pressupõe o uso de programa informático. Diante disso, a primeira conclusão seria a de que legislar sobre inteligência artificial insere-se na competência para legislar sobre informática.

Ocorre, contudo, que a repartição material de competências na CF se mostra mais complexa e relativizável, não apresentando uma delimitação tão precisa.

Isso acontece devido às intersecções e transversalidades das matérias que são objeto de proteção e regulação estatal confiadas a todos os entes federativos, as quais não raro apresentam pontos de sobreposição, interferência ou conflito.

tenha se popularizado no francês, a prioridade no uso do termo é atribuída ao alemão Karl Steinbuch, pelo pioneirismo na publicação do periódico *Informatik: automatische Informationsverarbeitung* (informática: processamento automático de informação). (PIMENTEL 2000, p. 29).

⁹ É que na versão dada por Vicentini, o sufixo “ics” teria sido adicionado à base do termo *Information*, no inglês, a exemplo de outros processos de derivação gramatical imprópria por sufixação, como *cybernetics*, *mathematics* etc., e não pela justaposição de informação e automática.

Por isso, é necessário em primeiro lugar uma taxinomia dos tipos de intervenção estatal, em que se constatem matérias ou âmbito de proteção ou regulação com ordem prioritária.

Um exemplo muito claro disso ocorre com a proteção do consumidor. O uso de inteligência artificial em uma relação de consumo não deve ser regulada exclusivamente pela União, já que a proteção ao consumidor é competência concorrente entre União, Estados e Distrito Federal (art. 24, V e VIII).

Também no caso de matérias de competência comum de todos os entes federativos, dispostas no art. 23 da CF, seria razoável admitir a exclusividade de intervenção estatal apenas pelo ente federativo União?

Assim não nos parece. No uso da competência privativa para legislar sobre informática, compete à União disciplinar aspectos relacionados ao processamento automático da informação em suas bases gerais, quando o bem jurídico protegido ou matéria tutelada não esteja inserida em outro âmbito material.

Além disso, o próprio conceito de inteligência artificial pode abrir espaço para a competência residual dos Estados. Se o foco for na decisão tomada automaticamente por máquinas, e não no processamento da informação, pode-se entender que inexistente competência da União para disciplinar a matéria, sendo a hipótese de competência residual.

Isso não impede que a União busque, com uma regulação, a proteção de direitos fundamentais ou discipline outras matérias que possam ser impactadas de forma significativa pelo uso dessa tecnologia.

Desse modo, para saber se o uso da inteligência artificial pode ser regulada pelos Estados, Distrito Federal e Municípios, devem ser seguidas as seguintes etapas:

1 - Identificar o âmbito material de intervenção: informática, direitos fundamentais, economia, direito civil, direito do consumidor, direito penal etc.;

2 - definir o escopo da regulação: proteção de direitos, atividade de fomento, aumento da segurança jurídica, uso pelo poder público ou adoção interna pela Administração, entre outros;

3 – constatar se a matéria regulada, considerado o escopo da regulação, é de competência privativa da União, de outro ente federativo, ou se insere no rol das competências comuns ou concorrentes.

As etapas 1 e 2 devem ser interpretadas em conjunto, pois as competências sobre determinada matéria podem ser diversas a depender da finalidade da regulação. Assim, criar regras gerais sobre consumo poderá atrair a competência privativa da União, mas a proteção do meio ambiente, da saúde ou das pessoas com deficiência, potencialmente impactadas pelo uso da inteligência artificial em uma certa relação de consumo, poderá justificar a atração da competência comum dos demais entes federativos.

Competência dos Municípios para regular a Inteligência Artificial

Considerando-se as etapas acima sugeridas, poderão os municípios legislar sobre inteligência artificial, nas seguintes hipóteses:

Predomínio do interesse local

A regra de ouro de repartição de competências preconstituída nos Estados federativos é o predomínio do interesse.

No Brasil, tal critério foi expressamente reconhecido na positivação de competência exclusiva dos Municípios para dispor sobre interesse local (art. 30, I, CF), posto que tal critério seja ainda sopesado com outros princípios federativos, como a necessidade de proteção suficiente e otimização de interesses dos demais entes em temas que possam ensejar indesejáveis competições interfederativas.

O conceito de *interesse local* não é preciso. Nas Constituições anteriores, empregava o constituinte a expressão *peculiar interesse*,¹⁰ cujo alcance sempre foi objeto de discussão doutrinária. Grande parte

¹⁰ Assim dispunha a Constituição brasileira de 1891: “Art. 68 - Os Estados organizar-se-ão de forma que fique assegurada a autonomia dos Municípios em tudo quanto respeite ao seu peculiar interesse.” A mesma expressão foi adotada nas Constituições de 1934 (Art. 13), de 1937 (art. 18), de 1946 (art. 28), de 1967 (art. 16), de 1969 (art. 15).

da controvérsia envolvia a discussão em torno da exclusividade do interesse municipal. (QUINTILIANO 2013, p. 247).

Sem embargo, a tese de que a expressão peculiar interesse (ou - agora – interesse local) não se confunde com interesse exclusivo parece prevalecer, especialmente devido ao fato de que muitos temas de interesse local não são propriamente exclusivos do Município, especialmente em questões pertencentes ao regime do federalismo cooperativo, como saúde, meio ambiente, segurança pública etc. A esse respeito, argumenta Roque Antônio Carraza:

Uma creche que o Município mantém é de seu interesse peculiar, porém não exclusivo, porque a proteção à infância importa, também, ao Estado e à União. Ainda a respeito, estabelece a Lei Magna ser a educação um ‘direito de todos e dever do Estado e da família’. (CARRAZA 2002, p. 158).

Melhor, então, é falar em interesse predominante. Será do Município a competência para legislar sobre assuntos de interesse local, desde que o interesse seja também exclusivo do Município, ou a fim de garantir maior proteção a um bem ou direito sem imposição de ônus desproporcional ao Estado ou à União, observando-se o espírito da lealdade federativa e do federalismo cooperativo.

Suplementação da legislação do estado ou da União

Além das competências consideradas próprias para dispor sobre assuntos de interesse local, em prestígio não apenas à sua autonomia, mas igualmente por força do princípio da subsidiariedade, os Municípios possuem competências inerentes à concepção cooperativa do federalismo trino brasileiro.

Desse modo, também podem os Municípios complementar a legislação federal e estadual (art. 30, II, CF).

Sobre a natureza legislativa ou material, há certa controvérsia doutrinária. Manoel G. Ferreira Filho (2000, p. 219), *v.g.*, entende que a referida suplementação consistiria, na verdade, na regulamentação da legislação federal e estadual, para sua execução. Seria, portanto, competência administrativa. Já Fernanda D. Menezes

de Almeida entende que o dispositivo encerra as duas competências, administrativa e legislativa (2005, p. 140).

Tal posição parece a mais correta, uma vez que se mostra mais coerente com o federalismo cooperativo e com as disposições constitucionais que outorgam ao poder público de qualquer esfera competências comuns para a proteção, fomento e garantia de acesso à fruição de direitos constitucionalmente assegurados.

Resta também saber quando tal suplementação é admitida. Fernanda D. Menezes de Almeida entende que tal suplementação apenas se justifica nas matérias de interesse local. (Ibidem).

Outra dificuldade é determinar quando "cabe" tal suplementação. A Constituição federal, nesse aspecto, não é tão clara, como quando faz em relação às competências concorrentes dos Estados.

Em primeiro lugar, devem-se distinguir as competências administrativas previstas no art. 23 da CF-88 que correspondam a competências legislativas previstas no art. 24 da CF-88 daquelas que não correspondam a uma competência legislativa específica. Na primeira hipótese, deve-se utilizar a mesma regra aplicável à competência suplementar e à supletiva dos Estados. Assim, quatro situações podem ocorrer: (1) não existem normas gerais; (2) existem normas gerais apenas da União; (3) existem normas gerais apenas dos Estados; ou (4) existem normas gerais da União, suplementadas pelos Estados. Seria possível o exercício de competência legislativa por parte dos Municípios nas quatro hipóteses? A resposta é afirmativa.¹¹

No caso de competências comuns que não encontram norma correlata no rol do art. 24, podem os Municípios atuar no campo da inteligência artificial, promovendo o fomento ou a proteção mais eficiente dos direitos envolvidos, observadas as demais limitações e princípios constitucionais, especialmente ligados ao pacto federativo brasileiro.

¹¹ Enfrentamos essa questão no trabalho "Autonomia federativa: delimitação no direito constitucional brasileiro". (QUINTILIANO, 2013, p. 248).

Rol exemplificativo de competências municipais para regulação direta ou reflexa do uso de IA

Diante das considerações acima, podem-se vislumbrar alguns exemplos de possíveis intervenções dos Municípios, materiais ou normativas, no tema da IA, como os propostos a seguir.

Transporte municipal

a) Competência plena (interesse local) – autorização para uso de aplicativos de transporte individual privado de âmbito municipal que utilizem IA; adoção ou não de aplicativos de transporte individual (táxi) ou coletivo (ônibus urbano, trem, metro etc.) que utilizem IA; uso da tecnologia no interior dos veículos destinados a tais serviços; permissão para uso de veículos autônomos.

b) Competência comum – proteção aos usuários de aplicativos de transporte ou aos motoristas mais eficiente que a conferida pelo respectivo Estado ou União, ou fomentadora da efetividade de direitos fundamentais, com a observância do princípio da razoabilidade entre outros.

c) Competência concorrente – detalhamento, complemento ou adaptação de legislação federal ou estadual para aplicação no âmbito local ou municipal.

Segurança pública

a) Competência plena (interesse local) – adoção ou não de sistemas de reconhecimento facial pelos órgãos municipais, em especial pela Guarda Municipal.

b) Competência comum – previsão de limites à adoção de sistemas de reconhecimento facial ou outra tecnologia com potencial de limitação da privacidade ou intimidade em espaços coletivos, públicos ou privados, que garantam proteção mais eficiente àquelas eventualmente ditadas pelo Estado ou pela União.

c) Competência concorrente - detalhamento, complemento ou adaptação de legislação federal ou estadual relativa ao uso ou seus efeitos de sistemas de inteligência artificial na área de segurança

pública ou à proteção de bens ou direitos que impliquem especialmente a atuação preventiva na área de segurança pública, para aplicação no âmbito local ou municipal.

Saúde

a) Competência plena (interesse local) - adoção ou não de recursos de inteligência artificial na execução de serviços públicos municipais de saúde.

b) Competência comum – previsão de limites para o uso de sistemas de inteligência artificial que causem prejuízo à saúde física ou mental no âmbito do Município, observados os demais princípios de repartição de competências;

c) Competência concorrente - detalhamento, complemento ou adaptação de legislação federal ou estadual relativa ao uso de sistemas de inteligência artificial na área da saúde, em suas diversas dimensões, como organização, serviços, medidas de proteção à saúde física e mental, para aplicação no âmbito local ou municipal.

Meio ambiente, fauna ou flora

a) Competência plena (interesse local) - adoção ou não de recursos de inteligência artificial para execução de suas competências materiais com impacto ambiental ou que afetem, de algum modo, o meio ambiente local ou a disciplina local sobre fauna ou flora; implementação de políticas públicas pautadas na concepção das *smart cities*.

b) Competência comum – limitação do uso de sistemas de inteligência artificial que causem prejuízo ao meio ambiente no âmbito do Município; disciplina sobre o uso de inteligência artificial para proteção da fauna e da flora, bem como eventuais limitações em sua utilização privada ou pública, quando potencialmente ostentarem riscos, observados os demais princípios de repartição de competências.

c) Competência concorrente - detalhamento, complemento ou adaptação de legislação federal ou estadual relativa ao uso de

sistemas de inteligência artificial na disciplina ou proteção do meio ambiente, da fauna ou da flora.

Proteção ao consumidor

a) Competência plena (interesse local) - adoção ou não de recursos de inteligência artificial na organização ou execução de serviços públicos municipais com aplicação subsidiária do Código de Defesa do Consumidor, nos termos da Lei federal 13.460/2017; adoção ou não, ou disciplina do uso, de sistemas de inteligência artificial para consecução das atividades de proteção ao consumidor, como o Procon e sistemas de solução *online* de controvérsias (*Online Dispute Resolution – ODR*).

b) Competência comum – disciplina sobre limites para o uso de sistemas de inteligência artificial nas relações de consumo no âmbito municipal, com vistas à proteção mais eficiente do consumidor, que não implique ônus excessivo e que não interfira de forma não razoável na disciplina estadual ou federal sobre a prestação desses serviços, especialmente quando objeto de autorização, concessão ou permissão de outro ente público.

c) Competência concorrente - detalhamento, complemento ou adaptação de legislação federal ou estadual de proteção ao consumidor, relativa ao uso de sistemas de inteligência artificial na relação de consumo, ou relativas ao Sistema Nacional de Defesa do Consumidor.

Educação, pesquisa, esportes, cultura e lazer

a) Competência plena (interesse local) - adoção ou não de recursos de inteligência artificial na organização ou execução de serviços públicos municipais de educação; fomento para utilização de sistemas de inteligência artificial nas áreas de educação, pesquisa, esportes, cultura e lazer; criação de espaços de estímulos à inovação nessas áreas (*sandbox* regulatório, contratos administrativos de impacto social ou em inovação, entre outras iniciativas).

b) Competência comum – uso de sistemas de inteligência artificial para proporcionar meios de acesso à cultura, à educação, à

ciência, à tecnologia, à pesquisa e à inovação (art. 23, V) ou para política de educação para a segurança no trânsito (art. 23, XII).

c) Competência concorrente - detalhamento, complemento ou adaptação de legislação federal ou estadual de proteção ao consumidor, relativa ao uso de sistemas de inteligência artificial para proteção e imputação de responsabilidade por danos ao patrimônio histórico e cultural.

Competência dos Estados-membros para regulação da Inteligência Artificial

Os Estados-membros, no Brasil, possuem competências exclusivas, implícitas ou explícitas, além de competências residuais, comuns, supletivas ou complementares.

Segundo o §1º do artigo 25 da Constituição Federal, são reservadas aos Estados todas as competências que não lhes sejam vedadas pela Constituição Federal.

A redação não é das mais felizes,¹² mas busca atribuir aos Estados competências residuais, ou seja, reconhecer aos Estados a competência para tratar de todas as matérias que não sejam atribuídas à União ou aos Municípios, explícita ou implicitamente.

Ocorre que, ao contrário de outros Estados federativos, onde a competência residual se mostra relevante, no Brasil, ela é extremamente limitada, tendo em vista que quase todas as competências estatais são expressa ou implicitamente atribuídas à União ou aos Municípios.

É necessário algum esforço para indicar exemplos de competências residuais. Alguns doutrinadores mencionam as competências administrativas, organizacionais e tributário-financeiras, como hipótese.¹³

¹² Na Constituição dos EUA, por exemplo, as competências residuais dos Estados-membros foram detalhadas na Emenda X, dispondo que “os poderes não delegados aos Estados Unidos pela Constituição, nem por ela negados aos Estados, são reservados aos Estados ou ao povo”.

¹³ Também reconhece o STF que questões ligadas ao direito administrativo inserem-se na competência residual de que trata o art. 25, §1º. Cf. ADI 1.279 MC, rel. min. Maurício Corrêa, j. 27-9-1995, P, DJ de 15-12-1995.

A par dessas competências implícitas, podem-se vislumbrar outras áreas, também reconhecidas pelo Supremo Tribunal Federal, como o turismo, a exploração de loterias e o transporte intermunicipal ou metropolitano.¹⁴

No julgamento da ADPF 492, ADPF 493 e ADI 4.986, o STF também afastou o entendimento de que as competências privativas legislativas da União previstas no artigo 22 da CF restringiriam competências materiais dos Estados, garantidas pelas competências residuais reconhecidas pelo §1º do art. 25 da CF. No caso, em que se discutia a possibilidade de exploração de loterias pelos Estados, entendeu a Corte que legislar sobre “consórcios e sorteios” não impede que os Estados prestem o serviço público de “sorteios”, na forma de loterias.

Assim, no que diz respeito à inteligência artificial, não obstante a previsão de que compete à União legislar sobre informática, possuem os Estados-membros a possibilidade de adoção de sistemas de inteligência artificial na prestação de serviços públicos.

Não obstante, é possível o entendimento de que legislar sobre inteligência artificial não seja matéria reservada à União ou aos Municípios, tampouco vedada aos Estados. Isso ocorre porque o conceito de informática, como visto, não absorve necessariamente por completo o âmbito regulatório em torno do uso de sistemas de IA. É claro que a disposição não pode afetar matérias de competência de outros entes federativos, como Direito Civil, Penal e a própria competência para dispor sobre informática etc..

Como exemplo, podemos citar uma autorização para utilização de carros autônomos ou até mesmo de prestação de serviços públicos por robôs. Disciplinar o uso de drones autônomos, *v.g.*, está inserido na competência para legislar sobre informática? E relações contratuais (talvez futuramente possíveis) entre seres humanos e inteligência artificial, como contratos, “casamentos”, ou

¹⁴ No RE 201.865 (rel. min. Carlos Velloso, j. 31-8-2005, P, DJ de 4-2-2005.) entendeu o STF que os Estados possuem competência para disciplinar o fretamento de ônibus para transporte com finalidade turística, à luz do art. 25, §1º, da CF. Cf. tb. ADI 2.349 (rel. min. Eros Grau, j. 31-8-2005, P, DJ de 14-10-2005). Sobre a loterias, cf. o recente entendimento adotado no julgamento da ADPF 492, ADPF 493 e ADI 4.986 (rel. min. Gilmar Mendes, j. 30-9-2020, P, DJE de 15-12-2020).

mesmo relações de ofensas, assédio ou quaisquer práticas consideradas preconceituosas dirigidas a robôs, devem ser interpretadas como inseridas na competência privativa da União para legislar sobre informática, ou se trata de competência residual dos Estados? Cabe lembrar que o direito civil regula relações entre particulares. A única relação entre pessoas e coisas regulada pelo Direito Civil diz respeito à propriedade ou posse, não a relações outras, como as de cunho afetivo, obrigacional ou de sucessão patrimonial, por exemplo.¹⁵

A par dessa discussão, cabe reconhecer que os Estados também possuem competências próprias, comuns e concorrentes, podendo disciplinar o uso da inteligência artificial na consecução dessas competências, a exemplo das hipóteses já descritas no item referente às competências municipais.

Cabe destacar que, em matéria de competência concorrente, as competências da União se limitam à edição de “normas gerais”.

O texto constitucional não define o que são normas gerais, nem se o adjetivo “gerais” se refere à extensão material ou subjetiva da norma. Para o STF e maior parte da doutrina, contudo, normas gerais são aquelas que fixam critérios, diretrizes, sendo dotadas de um certo grau de generalidade e abstração, que não esgote o assunto que será objeto de aplicação pelos Estados e Municípios.¹⁶ Estes precisam de certo espaço para aplicação das normas gerais e para sua adaptação. No julgamento da ADI 927, *v.g.*, também entendeu o STF que as normas gerais não podem limitar de forma considerável a autonomia dos Estados.

Em matéria de licitações, ainda, poderão Estados e Municípios disciplinar o uso da inteligência artificial em seus procedimentos, considerando que, segundo a Lei federal 14.133/2021, as licitações devem ser realizadas preferencialmente sob a forma eletrônica, devendo-se, ainda, recorrer à tecnologia para registros diversos, organização, disponibilização de dados etc.

¹⁵ Seria o caso de disciplinar “testamentos” deixados para robôs, sempre, é claro, *de lege ferenda*, ou ainda o reconhecimento de personalidade jurídica para sistemas de inteligência artificial, o que dependeria, também, de alteração no Código Civil.

¹⁶ Nesse sentido, cf. ADI 927 MC, Relator(a): CARLOS VELLOSO, Tribunal Pleno, julgado em 03/11/1993, DJ 11-11-1994.

Outra situação se refere à inteligência artificial como objeto da licitação, especialmente na modalidade especial criada pela Lei Complementar federal 182/2021, que institui o marco legal das *startups* e do empreendedorismo inovador.

A referida norma ainda apresenta pouca casuística, dada sua recentidade, mas, além de abrir competências para os Estados e Municípios na disciplina da aquisição de inovação tecnológica, onde se insere a inteligência artificial, também apresenta disposições que podem ser afastadas pelo Judiciário, por extrapolarem a competência da União para edição de normas gerais. É o caso da norma contida no §3º do art. 13 da referida lei, que exige um professor de instituição pública de educação superior para a participação na comissão especial que avaliará e julgará a contratação. Trata-se de detalhismo que avança no espaço que deve ser deixado a cada ente federativo, a par de fugir da sistemática do diálogo competitivo previsto na Lei 14.133/2021, modalidade que também pode ser empregada para aquisição de sistemas de inteligência artificial pela Administração Pública.¹⁷

Desse modo, além de sua competência residual, os Estados podem dispor sobre IA em serviços públicos de saúde, educação e segurança pública, entre outros, inseridos no âmbito de sua competência concorrente, ou mesmo na proteção dos bens jurídicos tutelados e descritos no art. 23 da CF, observando-se sempre os princípios e limitações impostas pela Constituição Federal (incluindo-se disposições contidas em tratados internacionais) ou por normas infraconstitucionais federais de caráter nacional.

¹⁷ A esse respeito, cabe registrar que a Prefeitura de São Paulo iniciou processo licitatório para aquisição de sistema de videovigilância urbana com inteligência artificial (*smart sampa*), na modalidade pregão, seguindo, ainda, as disposições da Lei federal n. 8.666/1993. Cf. edital em: https://www.bec.sp.gov.br/bec_pregao_ui/edital/becprp16001.aspx?VXMvviwwpHOiMQoeJWnSo8%2bctF4SMGPAiXN%2bfidhuMpKvl%2bYgqRzTMmgoPahg%2bs. Acesso em 10 set. 2023.

Considerações finais

A regulação do uso de sistemas de inteligência artificial deve considerar as progressivas relativizações dos conceitos clássicos de soberania e territorialidade, bem como os aspectos de interconstitucionalidade ou transconstitucionalidade, que passam a caracterizar os sistemas jurídicos contemporâneos.

Em outras palavras, a adoção de um regramento por um Estado soberano ou por um ente federado autônomo integrante de uma Federação não pode, a despeito de seu poder e direito de autodeterminação, fechar os olhos às tendências globais e a certa exigência de uniformidade de tratamento, o que pode representar para a entidade política isolamento, especialmente em termos econômicos.

Por outro lado, os Estados soberanos ou autônomos possuem a responsabilidade de atender aos fins perseguidos por seu povo, o que inclui a preservação de sua autodeterminação e o consequente direito de se opor a uma tendência global - ou transconstitucional - que vulnere seus princípios ou direitos fundamentais.

Assim, há que se reconhecer que, em sistemas de constitucionalismo multinível, opera o princípio do nível mais elevado de proteção dos direitos fundamentais, independentemente do sistema de repartição de competências adotado. No Brasil, tal princípio encontra amparo na interpretação do artigo 23 da CF, além de outras previsões esparsas na CF, sobretudo no Título VIII (Da Ordem Social – arts. 193 a 232 da CF).

No sistema brasileiro, que adota a forma federativa de Estado desde a Proclamação da República, em 1889 (Decreto n. 1, de 15 de novembro), Estados, Municípios e Distrito Federal possuem autonomia para adotar regulações sobre o uso da inteligência artificial, observando-se o sistema de repartição de competências desenhado pela Constituição de 1988.

Embora extremamente centralizado na União, o federalismo brasileiro reconhece aos entes subnacionais autonomia e

competências que reclamam um espaço de conformação para decidir sobre fomento, uso e proteção em face dessas novas tecnologias, mormente quando aplicadas em esferas materiais de sua competência exclusiva, comum ou concorrente.

A globalização, nacionalização ou regionalização são valores válidos e tendências inexoráveis na novel Sociedade da Informação. Eles não podem, contudo, subverter um dos princípios mais caros – senão o próprio núcleo – de qualquer sociedade política - seu direito à autodeterminação e à consequente preservação de sua história, cultura e valores. Que a conveniência e os apelos da uniformidade não ofusquem a riqueza e a defesa da diversidade.

Referências Bibliográficas

ALMEIDA, Fernanda Dias Menezes de (2005). **Competências na constituição de 1988**. 3. ed. São Paulo: Atlas.

BRASIL. Senado Federal. (2022). Comissão de juristas responsável por subsidiar elaboração de substitutivo sobre Inteligência Artificial no Brasil. **Relatório final**. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento?dm=9221643&ts=1670509473310&disposition=inline>.

Acesso em: 21 set. 2023.

CARRAZA, Roque Antônio. (2002). **Curso de direito constitucional tributário**. 19. ed. São Paulo: Malheiros.

ESTADOS UNIDOS DA AMÉRICA. (2019). **Executive Order 13859 of February 11, 2019**. Fev. 2019. Disponível em: <https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-americanleadership-in-artificial-intelligence>. Acesso em 20 set. 2023.

_____. (2020). Escritório de Administração e Orçamento do Escritório Executivo do Presidente dos EUA. *Memorandum for the Heads of Executive Departments and Agencies M-21-06*, nov. 2020. Disponível em: <https://www.whitehouse.gov/wp-content/uploads/2020/11/M-21-06.pdf>. Acesso em 15 set. 2023.

_____. (2023). *The National AI Advisory Committee – NAIAC*. Disponível em: <https://www.ai.gov/naiac/>. Acesso em 13 jun. 2023.

FERREIRA, M. G. (Fº) (2000) **Comentários à constituição brasileira de 1988**. 3. ed. São Paulo: Saraiva.

HUANG, S. *et al.* (2023) Translation: Measures for the Management of Generative Artificial Intelligence Services (Draft for Comment): Novel rules

about training data and accuracy of generated media circulated for comment. DIGICHINA. Stanford University. Apr. 2023. Disponível em: <https://digichina.stanford.edu/work/translation-measures-for-the-management-of-generative-artificial-intelligence-services-draft-for-comment-april-2023/>. Acesso em 21/09/2023.

PIMENTEL, A.F. (2000) **O direito cibernético**: um enfoque teórico e lógico-aplicativo. Rio de Janeiro: Renovar, 2000.

QUINTILIANO, L.D. (2012). **Autonomia federativa**: delimitação no direito constitucional brasileiro. Tese (Doutorado em Direito do Estado) - Faculdade de Direito, Universidade de São Paulo, São Paulo. Disponível em: <https://www.teses.usp.br/teses/disponiveis/2/2134/tde-26082013-162030/pt-br.php>. Acesso em: 2023-09-23.

UNIÃO EUROPEIA (2017). Resolução do Parlamento Europeu, com recomendações à Comissão de Direito Civil sobre Robótica. Disponível em: https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_PT.html. Acesso em: 15 ago. 2022.

_____. (2023). **Proposta de Regulamento da Inteligência Artificial**: Primeira Leitura. Disponível em: https://www.europarl.europa.eu/doceo/document/TA-9-2023-06-14_EN.html. Acesso em: 20 set. 2023.

VICENTINI, A. (1968) Informática. *Cadernos BAD* (Portugual) 3. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/67994>. Acesso em: 21 set. 2023.

Autor

Leonardo David Quintiliano

Especialista em Direito Digital pela Fundação Superior do Ministério Público do Rio Grande do Sul. Pesquisador Especialista em Direito Constitucional e Mestre em Ciências Jurídico-Políticas pela Faculdade de Direito da Universidade de Lisboa. Doutor em Direito do Estado pela Universidade de São Paulo. Professor de Direito Constitucional e Administrativo da Universidade Ibirapuera, da Escola Paulista de Direito e da Veni Creator Christian University (EUA). Associado Fundador do Instituto Avançado de Proteção de Dados (IAPD). Pesquisador do Grupo de Pesquisa em Direito, Ética e Inteligência Artificial da Faculdade de Direito da USP (Ribeirão Preto).

ISO E INTELIGÊNCIA ARTIFICIAL: guias e recomendações

João Paulo Candia Veiga
Laura Simões Camargo

Introdução

No momento em que o mundo discute a regulação da Inteligência Artificial (IA), se teremos uma governança global com a liderança de governos e Organizações Internacionais, ou se as empresas desenvolvedoras de aplicativos e sistemas de IA continuarão a dar as cartas, é muito bem-vinda a normatização de regras e normas para disciplinar o uso de tecnologias pela ISO e IEC. Na realidade, a normatização da IA através de padrões ou *standards* vem ocupar uma lacuna na medida em que os governos não conseguem chegar a um consenso para regulamentar esta tecnologia. Há muitas iniciativas que não convergem.

A Inteligência Artificial é, na realidade, uma designação genérica para um conjunto de tecnologias de propósito geral com muita assimetria de poder entre desenvolvedores, usuários, governos e organizações internacionais. Já a *International Standardization Organization (ISO)* e a *International Electrotechnical Commission (IEC)* são organizações internacionais de normatização, a primeira de propósito geral, privada, e a segunda dedicada à indústria, de caráter intergovernamental. Ambas criaram um sub-grupo de discussão para padrões de IA, dentro do *Joint Technical Committee (JTC)*. Por fim, a Associação Brasileira de Normas Técnicas (ABNT) é o Foro Nacional de Normalização dedicado à discussão de padrões técnicos desde a sua fundação, em 28 de setembro de 1940. É ela que representa o Brasil no *board* da ISO na Suíça, sendo a entidade privada responsável por

aprovar os padrões de IA no Brasil. Para tal, o primeiro passo é traduzi-los para o português.

A ISO separa os padrões em dois tipos: aqueles que farão parte das recomendações, e os que serão considerados requisitos, ou seja, exigências para as empresas buscarem a certificação, através de auditoria de terceira parte, como aconteceu com as famílias de ISOs para a qualidade (ISO 9000), e para a gestão ambiental (ISO 14001). Lembrando que os padrões ISOs são direcionados ao mercado, ou seja, é o setor privado que decide fazer o *compliance* aos padrões de forma voluntária.

O primeiro padrão ISO para a IA no Brasil é o ISO/IEC 22989, dedicado aos conceitos de Inteligência Artificial, para uniformizar a terminologia aplicada em português. O trabalho de tradução envolve a escolha de termos e de um vocabulário apropriado que precisa ser bem entendido e aceito pelas empresas. Por exemplo, *accountability* foi traduzido como “responsabilização”; *accountable* foi traduzido como “responsabilizado”; e *compliance* não foi traduzido, sendo decidido pela manutenção do termo em inglês, um indicativo de quanto o conceito já foi internalizado pelo setor privado no Brasil.

Em seguida, foi traduzido o ISO/IEC 38507 a respeito da governança de Tecnologia da Informação (TI) aplicada à IA. Por fim, seguiu-se a tradução do ISO/IEC 23894, com orientações às empresas sobre a gestão de riscos que envolvem as aplicações e sistemas de IA. Os três padrões estão em consulta pública no País e estarão disponíveis para adesões a partir de 2024. O próximo passo é traduzir o padrão que será requisito para a certificação de Inteligência Artificial no Brasil, a norma que diz respeito ao sistema de gestão de IA, a de número ISO/IEC 42001 que, na realidade, reúne todas as normas de TI aplicadas à IA.

A regulação privada através de normas ISO/IEC vem em boa hora e deve ajudar os governos a regular a Inteligência Artificial. O sistema ISO/IEC funciona a partir de consultas aos operadores da norma técnica em cada país, através de suas respectivas associações de normatização, ou seja, no caso de IA, os principais *stakeholders* a avaliarem a implementação das regras são as empresas de TI, *software* e aplicativos que desenvolvem algoritmos.

Ao mesmo tempo, há movimentação de governos e da Organização das Nações Unidas (ONU) para prover incentivos à governança global de IA. Referências a um “processo de IA” foram lançadas pelo G-7 em Hiroshima (maio de 2023), e o próprio Secretário Geral da ONU, António Guterres, defendeu a criação de uma autoridade global de IA já que o tema exige um “olhar universal”. De fato, está em andamento a criação de um conselho consultivo de alto nível multissetorial para Inteligência Artificial que apresentará, até o início de 2024, um relatório sobre as opções para governança global. As opções em tela vão da criação de uma nova agência da ONU até programas setoriais para implementar sistemas de IA éticos, de proteção aos Direitos Humanos. Em junho de 2023, o Parlamento Europeu aprovou a chamada “Lei da IA”, até o momento, o instrumento jurídico mais consistente de uma regulação pública, semelhante ao Projeto de Lei 2338 do Senado Federal brasileiro, com a categorização de risco para IA - inaceitável, de alto risco e risco limitado.

Por que normatizar o uso da Inteligência Artificial?

A proposta de normatização da Inteligência Artificial surge em um contexto de discussão internacional sobre o uso indiscriminado da tecnologia levar a consequências sociais nefastas, como a reprodução de racismo (através de ferramentas de *Facial Recognition*) e o reforço de desigualdades socioeconômicas (através de algoritmos enviesados que não aprovavam crédito para pessoas de determinadas cor e classe, por exemplo).

Mas, para compreender o porquê da padronização do uso da Inteligência Artificial, é preciso, antes, entender o que é a IA, suas possibilidades e riscos. Existem diversas definições da tecnologia. Uma delas é a da OCDE (2019), segundo a qual a Inteligência Artificial pode ser definida como:

*A machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy.*¹⁸

Além disso, a IA engloba diferentes tipos de tecnologia, como o *Machine Learning* (ML), o *Deep Learning* (DL), o *Natural Language Processing* (NLP) e a *Generative AI*.

A aplicação da tecnologia sem responsabilização tem gerado casos antiéticos no mundo todo, uma vez que o viés humano é introduzido através da estruturação do banco de dados que alimenta a IA, que o reforça em sua aplicação. Além disso, existe uma dificuldade de identificação da imputação desse viés, dado que o processo ocorre na chamada *black box*, não permitindo que haja transparência e muito menos responsabilização pelas ações. Dessa forma, considerando que a IA é uma tecnologia que tem sido aplicada em diversos campos sociais (como a saúde, mercado de trabalho, educação e cultura) e para a realização de múltiplas tarefas (do reconhecimento em fotos nas redes sociais¹⁹ à análise de dados sísmicos para identificação subterrânea de petróleo²⁰, por exemplo), tornou-se necessária a criação e adoção de normas, seja pelo setor público, seja pelo privado.

Neste cenário são postas duas grandes questões: como normatizar o uso da IA sem que o desenvolvimento das tecnologias e do setor privado sejam constrangidos e como essa regulação acompanharia o rápido desenvolvimento da IA dada a burocracia e morosidade das máquinas estatais? A normatização surge, então, como uma alternativa do setor privado de autorregulação, permitindo mais flexibilidade às demandas mercadológicas e agindo com mais rapidez quando necessário. A criação de padrões por órgãos de

¹⁸ Nota editorial: “Um sistema baseado em máquinas que pode, para um determinado conjunto de objetivos definidos por humanos, fazer previsões, recomendações ou tomar decisões que influenciam ambientes reais ou virtuais. Os sistemas de IA são concebidos para funcionar com diferentes níveis de autonomia”.

¹⁹ Disponível em: <https://correiodoestado.com.br/tecnologia/facebook-agora-reconhece-o-rostro-de-pessoas-nas-fotos/91050/>. Acesso em: 28 set. 2023.

²⁰ Disponível em: <https://www.poder360.com.br/energia/shell-usara-tecnologia-ia-para-explorar-petroleo-em-alto-mar/>. Acesso em: 28 set. 2023.

normatização sinaliza aos governos o caminho a seguir com a regulação pública.

As Normas ISO/IEC de Inteligência Artificial

A ISO e a IEC criaram um comitê conjunto com objetivo de estabelecer padrões internacionais de IA. O padrão ISO/IEC 22989 estabelece a terminologia, um vocabulário próprio de IA com a descrição de conceitos. No âmbito da ABNT, os conceitos foram discutidos e o vocabulário foi consensuado a partir dos agentes econômicos que operam os sistemas e aplicações de IA. Na realidade, o vocabulário é aplicável e serve a todas as organizações, desde empresas privadas e estatais até agências governamentais, fundações e organizações sem fins lucrativos.

No caso do ISO/IEC 23894, a norma dá indicações acerca de como as organizações que implantam ou utilizam produtos, sistemas e serviços que utilizam aplicações de IA devem gerenciar riscos relacionados à tecnologia. No processo de implementação, as organizações devem promover a integração do gerenciamento de riscos relacionados à IA e descrever como a gestão de riscos é desenvolvida. Dessa forma, os órgãos diretivos de uma organização podem garantir seu uso efetivo, dar transparência às suas práticas de gerenciamento de risco, e tornar o uso de IA mais ético e aceitável na perspectiva dos *stakeholders*, em particular, e da sociedade, de maneira geral.

O ISO/IEC 38507 trata da governança de Inteligência Artificial dentro das organizações. A norma aborda a natureza e os mecanismos da IA para entender as implicações de governança de seu uso: quais são as oportunidades, riscos e responsabilidades adicionais que o uso da IA traz? A ênfase está na governança do uso de IA pela organização e não nas tecnologias que compõem qualquer sistema de IA. No entanto, essa governança requer uma compreensão das implicações das tecnologias. A IA abrange uma família de tecnologias que reúnem poder computacional, escalabilidade, rede, dispositivos e interfaces conectados, juntamente com grandes quantidades de dados, o que torna necessária uma estrutura de governança complexa e integrada. Existe sempre a possibilidade de que a direção e os controles

existentes não sejam apropriados para garantir os resultados necessários (e a mitigação do risco de consequências indesejáveis).

Por um lado, a utilização de sistemas de IA sem conhecimento ou consideração de potencial viés, erros ou danos, ou das implicações da incorporação da IA em sistemas complexos existentes ou mesmo o impacto da IA no mercado de trabalho, incluindo preocupações com a discriminação, os danos aos direitos humanos ou aos direitos fundamentais dos trabalhadores, a redundância devido à automatização ou à requalificação e a possível perda de conhecimento organizacional, mas também a alavancagem da IA para aumentar a criatividade humana, aumentar a qualidade do trabalho delegando tarefas repetitivas, triviais ou perigosas à um sistema de IA.

A norma ISO determina ainda que os encarregados pela organização que utilizam ferramentas de IA podem ser responsabilizados pelas más condutas da empresa nos casos em que diligência, cuidado, orientação, treinamento, supervisão e implementação inadequados dentro da estrutura de governança permitam o surgimento de problemas. Essa responsabilização pode ser assegurada pela direção da empresa ou imposta pelas partes interessadas ou por outros meios. Os membros do órgão diretivo podem enfrentar uma penalidade, remoção do cargo ou reparação legal. Dessa forma, a direção da empresa deve garantir que suas práticas sejam adequadas para os usos específicos aos quais a IA está sendo aplicada dentro da organização. Isso pode incluir a revisão e, quando necessário, o aprimoramento de ações, programas e políticas, estratégia, alocação de recursos, códigos de ética, manifestações de valores, propósito ou outros instrumentos relacionados ao uso da IA na organização. A direção da empresa tem a obrigação de garantir a supervisão responsável através de uma avaliação das aplicações de IA realizadas, uma verificação do seu valor para a organização e do apetite ao risco da empresa, e garantia de implementação, monitoramento e medição dos sistemas de IA utilizados.

A estrutura de governança deve prever ainda instrumentos de avaliação que considera diferentes elementos, por exemplo, os fatores internos e externos relacionados à organização, ameaças e oportunidades atuais e futuras, resultados alcançados, eficácia e eficiência dos mecanismos de governança em vigor e julgamentos

sobre decisões e opções tomadas. Por fim, relatórios periódicos devem demonstrar às partes interessadas que a utilização da IA está sendo efetivamente controlada pelos responsabilizados.

De acordo com a norma ISO, embora a empresa possa identificar facilmente os benefícios do uso da IA, convém que a direção da empresa entenda as restrições e obrigações que esse uso coloca para a organização. Para administrar adequadamente essas restrições e obrigações, recomenda-se, assim, que a estrutura de governança da organização preveja ações importantes como aumentar a supervisão de *compliance*. Sem a supervisão apropriada, o uso da IA pode automatizar processos e produzir resultados que promovam mudanças frequentes, o que pode ser de difícil explicabilidade ou pode entrar em conflito com outras ações, programas e políticas organizacionais.

Além disso, o escopo da automação dos sistemas de Inteligência Artificial deve ser rigorosamente supervisionado pelo órgão diretivo da empresa e implementado por pessoas devidamente autorizadas e qualificadas. A responsabilidade necessária deve ser mantida e as consequências do uso de tal sistema precisam ser testadas e propriamente compreendidas antes da implementação.

A estrutura de governança precisa ainda prever os impactos do uso de aplicações de IA sobre as partes interessadas. Embora algumas decisões tenham um impacto negativo sobre os atores envolvidos (por exemplo, recusar um empréstimo bancário a um cliente), a empresa precisa assegurar que tais impactos não sejam exacerbados pelo uso da IA. Pela norma ISO, no caso de requisitos legais ou obrigações para o uso de sistemas de IA, a estrutura de governança deve dar indicações precisas. É cada vez mais provável que a utilização de IA e soluções associadas, como o reconhecimento facial ou movimento automatizado de veículos, por exemplo, sejam contestadas e possam ser objeto de novos requisitos legais. A empresa precisa demonstrar de que forma essas obrigações são cumpridas e se estão em conformidade com os requisitos e devidamente explicadas, quando necessário.

O uso de IA implica em inovação, e sabe-se que a utilização de novas tecnologias é fundamental para a viabilidade e a saúde de muitas empresas. Nesse caso, a governança precisa incentivar a

inovação e não se contrapor a ela. Nem todo projeto que envolve IA será estrategicamente importante (por exemplo, alguns só reduzirão custos), mas, no geral, convém que o uso da tecnologia ajude a organização a atingir seus objetivos. O atingimento dos objetivos deve acontecer junto ao uso de IA alinhado à cultura e aos valores da empresa, bem como à ética corporativa.

No caso do ISO/IEC 42001, o objetivo central é orientar as empresas a respeito de como gerenciar seus sistemas de IA. Trata-se, portanto, de uma norma de gestão, assim como a família ISO 14001 é endereçada à gestão ambiental. Isso é importante, já que o uso de IA e ML (*machine learning*, um subconjunto da IA) implicam em transparência e explicabilidade dos sistemas de decisão automatizados, e no uso de saídas como a análise de dados de sistemas de ML, que são treinados a partir de dados, uma vez ou continuamente, e se adaptam às mudanças em suas entradas. Isso difere da programação processual tradicional, uma vez que um sistema de IA pode mudar seu comportamento ao longo de seu uso.

A ênfase da ISO/IEC 42001 está na integração de um sistema de gestão de IA com as estruturas existentes da organização, esse é o principal objetivo da norma. Dessa forma, um sistema de gestão é definido como um conjunto de elementos integrados de ações, programas e políticas corporativas, com objetivos bem definidos. Para capturar os componentes de um sistema de gestão, a norma opera em um determinado nível de generalidade, que deve servir a qualquer empresa, e traz, formalmente, a interação com outras normas ISO/IEC para obter a correta perspectiva do uso de aplicações de IA.

A norma ISO/IEC 42001 traz quatro anexos importantes. Uma Inteligência Artificial confiável é mencionada no anexo A como parte do guia de gerenciamento para o desenvolvimento de sistemas de IA. O anexo B trata das diretrizes de implementação para controles de IA com a indicação objetiva a medidas específicas relacionadas à IA/ML. Por exemplo, é obrigatório que a documentação dos dados utilizados na empresa inclua as categorias utilizadas para o ML e o processo de rotulagem dos dados para formação e ensaio. No que diz respeito à avaliação do impacto dos sistemas de IA em grupos e indivíduos, a norma menciona várias áreas de confiabilidade, como justiça, transparência, explicabilidade, acessibilidade e segurança. Outras

áreas são indicadas pela norma, como o efeito sobre o meio ambiente, a potencial desinformação e possíveis problemas adversos de segurança e saúde.

O anexo B da norma prevê ainda uma justificativa para o desenvolvimento de um sistema de IA, com a respectiva explicação de quando e por que o sistema será usado, além de uma lista de métricas que devem ser adotadas para medir se o desempenho do sistema é compatível com esses objetivos. Uma questão que se coloca é se as métricas usadas para sistemas de *software* são suficientes para aqueles sistemas que contêm IA. O anexo B entende também que as opções de *design* devem ser documentadas, o que inclui detalhes sobre o método ML.

O anexo B também aborda os processos de gestão de dados. Eles abrangem orientações para sistemas de IA que incluem transparência, explicabilidade e dados de treinamento de amostra, além de orientações para a preparação de dados, e a exploração estatística dos dados. O uso responsável da IA também é abordado, tendo como objetivo a confiabilidade do sistema de IA em várias dimensões, incluindo justiça, responsabilidade, transparência, robustez, segurança, privacidade e acessibilidade.

O anexo C traz os objetivos da organização no que se refere às fontes de risco relacionados à IA. E o anexo D do ISO/IEC 42001 trata dos domínios e setores em que um sistema de IA pode ser utilizado. O problema em tela é a integração de padrões específicos do setor com o padrão geral de IA. Este anexo também trata da certificação, com uma abordagem comum às normas ISO/IEC de promoção de avaliação da conformidade a partir de auditoria de terceira parte.

De maneira geral, a norma ISO/IEC 42001 oferece às organizações a orientação para os padrões e processos necessários para que a implantação de IA construída sobre eles seja minimamente confiável, ao longo de suas várias dimensões. Na medida em que as organizações comecem a utilizar as normas ISO/IEC de IA, e a certificação voluntária for bem aceita pelo mercado, é possível que essa dinâmica influencie os governos no momento de definir uma regulação pública de IA. A interação regulatória público-privada de IA pode ajudar os países a avançarem na redução dos riscos e em maximizar os impactos positivos da IA.

Conclusão

A normatização da Inteligência Artificial pela *International Standardization Organization* (ISO) e a *International Electrotechnical Commission* (IEC), em nível internacional, e pela Associação Brasileira de Normas Técnicas (ABNT), em nível nacional, tem gerado diversas reflexões sobre a gestão de sistemas de IA, o gerenciamento de riscos do uso da tecnologia, governança, domínios e setores em que a IA pode ser usada, entre outros.

A criação dessas normas pode ser compreendida como uma resposta do setor privado em meio ao debate sobre a criação de padrões que disciplinam o uso da IA. Por um lado, enquanto a decisão estatal brasileira ainda não foi tomada e a discussão se complexifica ao tratar também de desenvolvimento econômico nacional e de Ciência, Tecnologia e Inovação (CT&I), o setor privado apresenta uma alternativa com maior rapidez e flexibilidade. Essas características das normas e padrões privados são especialmente úteis no caso de tecnologias que mudam rapidamente, como a IA. No entanto, por outro lado, a autorregulação privada abre espaço para a não responsabilização efetiva das consequências do uso da ferramenta, bem como tira o Estado de sua posição fiscalizadora e, quando necessário, punidora do setor privado. Desta forma, as normas ISO/IEC podem preencher uma lacuna importante entre a autorregulação sem disciplina, e as regras e normas de governos e organizações internacionais.

Referência

OECD's homepage on legal instruments:
<https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449>

Autores

João Paulo Candia Veiga

Cientista Político. Professor do Departamento de Ciência Política da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo (FFLCH-USP). Pesquisador do Centro de Estudos das Negociações Internacionais (CAENI). Pesquisador principal do Centro de Inteligência Artificial (C4AI).

Laura Simões Camargo

Bacharela em Ciências Sociais pela Universidade de São Paulo (USP). Pesquisadora do Center for Artificial Intelligence (C4AI) e da Fundação José Luiz Egydio Setúbal. Trabalhou na Secretaria Municipal de Direitos Humanos e Cidadania de São Paulo e no Centro Brasileiro de Análise e Planejamento (CEBRAP).

IA GENERATIVA:

passos essenciais para a implantação nas organizações

Fábio Correa Xavier

Introdução

A transformação digital emergiu como uma imperativa necessidade para organizações contemporâneas buscando prosperar em um ambiente em constante mutação. Nesse cenário, a Inteligência Artificial (IA) se destaca como uma disciplina essencial da ciência da computação, centrada na capacidade de criar sistemas e máquinas capazes de executar tarefas tradicionalmente reservadas para a inteligência humana. Dentro desse amplo domínio, a IA generativa assume um papel de destaque, permitindo que as organizações desbloqueiem um potencial criativo sem precedentes.

Este artigo explorará os principais passos a serem considerados ao implementar a IA generativa nas estratégias das organizações. Abordaremos a transformação digital e a IA generativa, detalhando como essa categoria de IA é capaz de não apenas compreender e responder a informações, mas também de criar conteúdo novo e inovador, imitando a criatividade humana.

Ao longo do artigo, examinaremos cada etapa do processo de implementação da IA generativa. Desde a autorização formal necessária para o uso de dados até a definição precisa dos problemas que essa tecnologia pode solucionar. Exploraremos a preparação dos dados, fundamental para garantir que as informações estejam disponíveis e formatadas adequadamente para treinar os modelos. Além disso, abordaremos o treinamento do modelo, um passo crítico

que capacita a IA generativa a desempenhar suas funções de maneira eficaz.

Finalmente, discutiremos a validação e a implantação desses modelos nas operações da organização, enfatizando a importância da melhoria contínua. Ao longo do texto, ilustraremos como a IA generativa pode revolucionar as operações e estratégias das organizações, proporcionando uma vantagem competitiva notável e ampliando a capacidade criativa em diversas áreas, desde a geração de conteúdo até a solução de problemas complexos.

Desvendando a Inteligência Artificial Generativa: conceitos fundamentais

A transformação digital se estabeleceu como um imperativo para as organizações modernas que desejam prosperar em um cenário de constante evolução. A Inteligência Artificial é um campo amplo da ciência da computação que se concentra em criar sistemas ou máquinas capazes de realizar tarefas que normalmente requerem inteligência humana. Isso inclui áreas como aprendizado de máquina, processamento de linguagem natural, visão computacional, robótica e muito mais.

Por sua vez, os Modelos de Linguagem são algoritmos de IA projetados para entender e gerar texto em linguagem natural. Os Grandes Modelos de Linguagem (*Large Language Models* - LLM) são uma categoria de modelos de linguagem que se destacam por sua escala, com milhões ou até bilhões de parâmetros. Eles são treinados em grandes quantidades de texto não rotulado e podem ser ajustados para tarefas específicas.

O GPT (*Generative Pre-trained Transformer*) é um tipo de *Large Language Model* (LLM) desenvolvido pela OpenAI. Ele é uma implementação avançada de inteligência artificial que utiliza a arquitetura *Transformer* para entender e gerar texto em linguagem natural de maneira coesa e contextual. O GPT se destacou por sua capacidade de realizar uma variedade de tarefas de processamento de linguagem natural de forma impressionante, sem a necessidade de treinamento específico para cada tarefa. Em vez disso, ele é pré-

treinado em grandes quantidades de texto não rotulado e pode ser ajustado (*finetuned*) para tarefas específicas.

A arquitetura *Transformer*, na qual o GPT é baseado, é composta por várias camadas de atenção, permitindo que o modelo compreenda a relação entre palavras em uma frase ou contexto mais amplo. Isso dá ao GPT a capacidade de gerar texto que parece escrito por um ser humano e de entender a semântica e o contexto de maneira profunda.

A Inteligência Artificial Generativa representa um novo paradigma nos sistemas de IA, permitindo que a IA demonstre e imite a criatividade humana, tornando-a cada vez mais útil em carreiras criativas e intensivas em conhecimento, como publicidade, publicação, entretenimento e desenvolvimento de software e codificação de aplicativos (Berg et al., 2023; Dwivedi et al., 2023). Os casos de uso do ChatGPT abrangem desde a redação e correção de materiais de marketing, planos de negócios, documentos legais e código de software, até a elaboração de respostas por e-mail, resumos de texto, tradução e identificação de anomalias, erros e defeitos em materiais textuais. Estudos recentes demonstraram que o desempenho do ChatGPT em algumas tarefas de trabalho do conhecimento supera a qualidade dos trabalhadores humanos (GILARDI *et al.* 2023).

Alguns exemplos de aplicações para a Inteligência Artificial Generativa seriam:

1. Geração de Texto Criativo: o GPT pode ser usado para gerar histórias, poesias, redações, entre outros tipos de texto criativo. Ele é capaz de criar conteúdo coeso e convincente com base nas informações e no estilo fornecidos.
2. Respostas a Perguntas: o GPT pode responder a perguntas com base no contexto fornecido. Por exemplo, pode responder a perguntas gerais de conhecimento, fornecer explicações detalhadas sobre um tópico ou até mesmo criar respostas para questões específicas.
3. Tradução Automática: o GPT pode ser usado para traduzir texto de um idioma para outro de maneira fluente e precisa.
4. Assistência em Escrita: o GPT pode ajudar na redação de e-mails, textos acadêmicos, relatórios e outros tipos de

documentos, fornecendo sugestões de palavras, frases e estrutura.

5. Atendimento ao Cliente Automatizado: empresas podem usar o GPT para criar *chatbots* que interagem com os clientes de forma natural, respondendo a perguntas e solucionando problemas comuns.

A OpenAI tem desenvolvido o GPT e atualmente temos as seguintes versões:

1. GPT-1: a primeira versão do GPT tinha 117 milhões de parâmetros e foi um marco inicial na demonstração das capacidades dos LLMs.
2. GPT-2: a segunda versão, com 1,5 bilhão de parâmetros, chamou muita atenção devido à sua capacidade de gerar texto extremamente coeso e fluente. Sua capacidade de produzir conteúdo fictício levou a preocupações sobre desinformação.
3. GPT-3: a versão mais utilizada, GPT-3, possui 175 bilhões de parâmetros e é considerada um dos maiores e mais poderosos modelos de linguagem já criados. Ele pode realizar uma ampla variedade de tarefas, desde conversas em linguagem natural até programação de computadores.
4. GPT-4: versão mais recente do GPT, lançada em março de 2023, tem uma melhor capacidade de compreensão e geração de texto, pois estima-se que ele possua em torno de 1 bilhão de parâmetros.

Essas versões do GPT têm demonstrado um grande progresso na compreensão e geração de linguagem natural, impulsionando avanços em uma variedade de campos, incluindo educação, comunicação, atendimento ao cliente e criação de conteúdo, dentre outros.

O LaMDA - *Language Model for Dialogue Applications* – é outra implementação de IA Generativa, desenvolvido pela Google AI, e projetado especificamente para aplicações de diálogo. Ele também utiliza a arquitetura *Transformer*, mas é ajustado (*finetuned*) para ter um melhor desempenho em interações de conversação, permitindo

uma compreensão mais profunda do contexto e nuances das conversas. O LaMDA foi treinado em um enorme conjunto de dados de texto e código. Ele é capaz de gerar texto, traduzir idiomas, escrever diferentes tipos de conteúdo criativo e responder a perguntas de forma informativa. O LaMDA tem 137B parâmetros, o que o torna um dos modelos de linguagem mais complexos do mundo. Ele é treinado em um enorme conjunto de dados de texto e código, o que lhe permite gerar texto que é factualmente correto e criativo.

Mas nem tudo são flores quando falamos de IA Generativa, como o próprio ChatGPT. Os principais problemas estão relacionadas à maneira como os grandes modelos de linguagem operam, pois eles fazem previsões complexas sobre o que provavelmente é a saída "correta" ou desejada com base nas instruções do usuário (Edelman e Abraham, 2023). Com o ChatGPT, isso pode resultar em respostas incorretas ou tendenciosas e em um fenômeno chamado "alucinação", sendo que muitos de nossos informantes observaram que não usariam o ChatGPT como a única fonte de informações factuais, mas sim como um "parceiro de diálogo" para geração de ideias.

É importante esclarecer esses conceitos tratados até aqui. A Inteligência Artificial é uma área da ciência da computação que se concentra em criar sistemas que podem realizar tarefas que normalmente requerem inteligência humana, como aprendizado, raciocínio e resolução de problemas. O *Deep learning* é uma subárea da IA que se baseia em redes neurais artificiais profundas para aprender e realizar tarefas complexas, como reconhecimento de voz e imagem. Os modelos de linguagem, por sua vez, são algoritmos de IA projetados para entender e gerar texto em linguagem natural. Eles são usados para tarefas como tradução automática e resumo de texto. Os Grandes Modelos de Linguagem são versões avançadas de modelos de linguagem que têm uma enorme quantidade de parâmetros e são treinadas em grandes quantidades de texto. O GPT e o LaMDA são exemplos de implementações de LLM. O GPT foi desenvolvido pela OpenAI e é capaz de entender e gerar texto em linguagem natural e pode ser usado em várias aplicações, como *chatbots* e geração de texto. O LaMDA é outro modelo de linguagem desenvolvido pela Google. Ele é projetado para entender e responder a consultas complexas em linguagem natural e é usado principalmente para

melhorar as respostas de pesquisa do Google. Por sua vez, a IA generativa refere-se a sistemas de IA que podem criar conteúdo, como texto, imagens e música, de forma autônoma. Modelos de linguagem como GPT e LaMDA são exemplos de IA generativa, pois podem gerar texto em linguagem natural de maneira coesa e contextual.

Em resumo, a IA é o campo amplo, o *deep learning* é uma técnica dentro da IA, os modelos de linguagem são uma aplicação específica dentro do *deep learning*, os *Large Language Models* como GPT e LaMDA são versões avançadas desses modelos de linguagem, e a IA Generativa refere-se a sistemas de IA capazes de gerar conteúdo de forma autônoma, onde os modelos de linguagem desempenham um papel importante nesse processo. A figura a seguir ilustra o relacionamento desses conceitos.

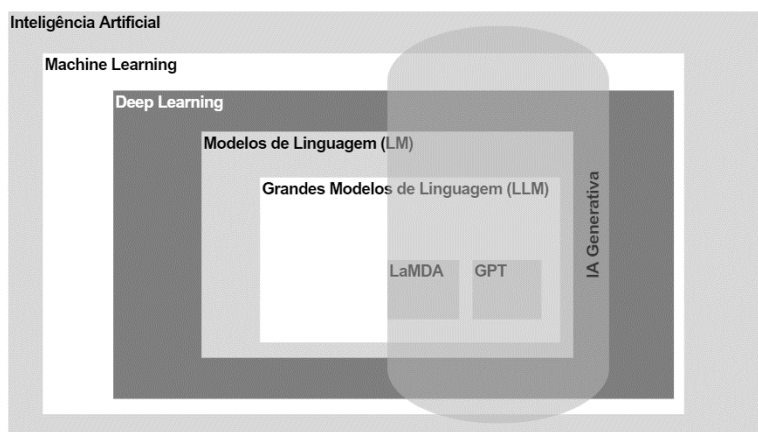


Figura 1 - IA e seus conceitos. Fonte:

Algumas oportunidades para aplicação da IA Generativa e grandes modelos Linguagem seriam:

1. Atendimento Aprimorado ao Cliente: esses modelos capacitam as empresas a oferecerem um atendimento ao cliente mais eficiente e personalizado. *Chatbots* baseados em linguagem natural podem responder a perguntas, solucionar problemas e

fornecer informações relevantes de maneira rápida e precisa, melhorando a experiência do cliente.

2. Automação de Processos: os modelos de Geração de Linguagem Natural podem automatizar uma variedade de tarefas, desde a geração de relatórios até a redação de conteúdo. Isso libera recursos internos e permite que os funcionários se concentrem em atividades de maior valor.
3. Tomada de Decisões baseada em informações: a análise de dados é essencial na transformação digital. Modelos desse tipo podem analisar grandes volumes de dados não estruturados, identificar padrões e fornecer *insights* valiosos para embasar decisões estratégicas.
4. Engajamento do Cliente: a criação de conteúdo envolvente é crucial para atrair e reter clientes. Os modelos de processamento de linguagem natural podem ajudar na geração de conteúdo relevante e cativante, desde postagens em redes sociais até campanhas de *marketing* personalizadas.

Desafios e oportunidades

Como acontece com qualquer tecnologia, há alguns desafios que precisam ser encarados quando da implementação desses modelos na transformação digital das empresas, especialmente quando se trata de uma inovação tão poderosa quanto a Inteligência Artificial Generativa. Esses desafios são multifacetados e demandam uma abordagem cuidadosa para garantir que a IA generativa seja aplicada de forma eficaz e ética em toda a organização. Dentre os desafios, os principais seriam:

1. Qualidade dos Dados: a eficácia do uso da IA Generativa depende da qualidade dos dados de entrada. Dados imprecisos ou desatualizados podem levar a resultados inadequados. As organizações precisam garantir a integridade e a relevância dos dados utilizados.
2. Treinamento e Aprendizado: os modelos de linguagem natural requerem treinamento extensivo e ajustes para se alinharem

às necessidades específicas da organização. Isso exige tempo e recursos significativos.

3. Privacidade e Segurança: o uso de dados sensíveis na implementação dos modelos levanta preocupações de privacidade e segurança. É essencial garantir a conformidade com regulamentações como, por exemplo, a Lei Geral de Proteção de Dados - LGPD.

É notório que a tecnologia de Inteligência Artificial Generativa, com sua capacidade de geração de linguagem natural, como o ChatGPT ou o Bard da Google, tem o poder de revolucionar a maneira como as organizações operam na era digital. Ao abraçar suas oportunidades e superar desafios, as organizações podem impulsionar a inovação, melhorar a eficiência e fornecer experiências excepcionais para clientes e funcionários.

Estratégias Eficientes para Implementar a IA Generativa nas Organizações

Está cada vez mais evidente que ferramentas de IA generativa, como o ChatGPT e o Bard, têm aproximado os seres humanos e a tecnologia de maneiras novas e mais intuitivas. Nesse sentido, utilizamos o termo "assistência algorítmica" para descrever como o ChatGPT já está transformando o trabalho de conhecimento nas empresas. O lado "algorítmico" destaca os processos sofisticados que estão no cerne da IA generativa, como o ChatGPT, que possibilitam a criação de resultados de valor com base na entrada do usuário. Ao escolher a palavra "assistência", enfatizamos que o ChatGPT ajuda os seres humanos a realizarem - ou evitarem - tarefas essenciais no trabalho de conhecimento, impactando assim os papéis de muitos funcionários na vida profissional contemporânea (DE CREMER *et al.* 2023).

No entanto, a implementação institucional da IA Generativa em uma organização é um processo que envolve uma série de etapas estratégicas. Além disso, é crucial contar com uma equipe altamente especializada e multidisciplinar, com conhecimento sólido em tecnologias de processamento de linguagem natural (*Natural*

Language Processing - NLP) e inteligência artificial (IA) (Jurafsky e Martin, 2020), bem como um profundo entendimento das considerações de privacidade de dados (Kroll et al., 2019) e requisitos do negócio.

Essa combinação de conhecimento técnico e domínio de negócios é crucial para o sucesso da implementação da IA Generativa em organizações, como sugerido por Hwang e Lee (2019) em seu estudo sobre a aplicação de IA nas estratégias empresariais. Portanto, a formação de equipes capacitadas e interdisciplinares é essencial para sua adoção de maneira eficaz.

Antes de se utilizar a IA Generativa na organização, é preciso buscar meios para se criar valor a partir dela. Para isso, as seguintes diretrizes podem ser seguidas para maximizar o potencial de criação de valor dessa tecnologia:

(i) Transformação funcional, repense processos: buscar uma transformação funcional ao invés de Casos de Uso, colocando o foco no redesenho de processos de negócios completos, em vez de se concentrar apenas em casos de uso específicos. O objetivo é obter uma transformação profunda e abrangente, que abarque toda a organização e não apenas partes isoladas. Isso envolve repensar como a IA generativa pode ser aplicada em toda a cadeia de valor da empresa.

(ii) Concentração estratégica, poucas oportunidades: concentrar-se em poucas oportunidades para construir uma massa crítica de mudança, não dividindo esforços em muitos projetos, pois é mais eficaz concentrar recursos em algumas oportunidades estratégicas. Isso ajuda a criar uma "massa crítica" de mudança, onde o impacto da IA generativa se torna mais evidente e significativo em um conjunto limitado de áreas.

(iii) Complementaridade IA e IA Generativa: avaliar o uso da IA tradicional e da IA Generativa, diferentes capacidades e oportunidades complementares, pois elas têm capacidades distintas. Elas não competem, mas sim se complementam em áreas específicas. Portanto, é importante entender as capacidades de ambas e identificar onde podem trabalhar juntas de forma sinérgica.

(iv) Vantagem competitiva inovadora: não se trata apenas de produtividade, pois a IA Generativa pode – e deve – criar vantagem

competitiva. As organizações podem usar a tecnologia para oferecer experiências de cliente inovadoras, *insights* valiosos e novas formas de inovação.

(v) Capacidades amplas da IA Generativa: além de gerar conteúdo, a IA Generativa possui capacidades mais amplas, como síntese de informações, design, resolução de problemas e controle de sistemas autônomos. Explorar todas essas capacidades é essencial para aproveitar todo o potencial da tecnologia.

(vi) Desafios na implementação: implementar a IA Generativa não é simples nem gratuito. Isso envolve desafios de engenharia, como projetar arquiteturas de dados adequadas e tomar decisões tecnológicas críticas ao escolher plataformas e modelos. Além disso, é necessário realizar testes rigorosos para garantir que o sistema funcione conforme o esperado.

(vii) Planejamento da força de trabalho: a implementação da IA Generativa tem impactos tanto quantitativos quanto qualitativos nas pessoas. Isso inclui a necessidade de planejar a força de trabalho para acomodar a tecnologia e, ao mesmo tempo, realizar um amplo programa de aprimoramento e reciclagem de habilidades.

(viii) Gerenciamento cuidadoso de riscos: é importante gerenciar cuidadosamente os riscos relacionados à criatividade e à diversidade de pensamento. O uso da IA Generativa pode influenciar a criatividade e a diversidade de ideias, portanto, é necessário um monitoramento rigoroso e medidas de mitigação para garantir que esses aspectos não sejam comprometidos.

Passos para implementação da IA generativa

A utilização da Inteligência Artificial Generativa oferece um vasto potencial para impulsionar a inovação e a eficiência nas organizações modernas. No entanto, essa jornada rumo à integração bem-sucedida da IA generativa em uma organização é repleta de desafios e etapas cruciais a serem seguidas. Neste contexto, é essencial compreender a importância de cada uma dessas etapas para obter resultados positivos. Ao seguir uma estrutura cuidadosamente delineada, as organizações podem evitar potenciais armadilhas legais, problemas de privacidade de dados e garantir que a IA generativa se

torne uma ferramenta eficaz para resolver problemas complexos e impulsionar a inovação.

De maneira sucinta, os passos para a utilização da IA generativa em uma organização seriam:

- Passo 1 - Autorização Formal: É crucial obter a autorização formal da organização para usar os dados em soluções de IA Generativa. A autorização formal deve ser explícita, incluindo detalhes sobre quais dados serão usados, como eles serão tratados e para quais fins serão utilizados. Se essa etapa não for seguida, a organização pode enfrentar sérias questões legais e de privacidade, além de possíveis violações das regulamentações de proteção de dados, como a LGPD no Brasil e a GDPR na União Europeia.
- Passo 2 – Delimitação do Problema: é importante definir claramente o problema que está sendo resolvido. Se o problema não for adequadamente delimitado, a organização pode desperdiçar recursos valiosos em análises de dados inúteis ou não relacionadas aos objetivos reais.
- Passo 3 - Preparação dos Dados: esta é uma etapa crítica, mencionada, pois a falta de preparação adequada dos dados pode levar a problemas como vazamento de dados pessoais, o que pode resultar em violações de privacidade e danos à reputação da organização. Além disso, não criptografar dados sensíveis pode expor informações confidenciais.
- Passo 4 - Treinamento do Modelo: Ao treinar o modelo selecionado, a organização deve seguir as boas práticas de ajuste fino (*fine tuning*) recomendadas pela desenvolvedora da solução escolhida. Ignorar essas práticas pode resultar em um modelo que não atende às necessidades específicas da organização e pode gerar respostas imprecisas ou inadequadas. A falta de treinamento apropriado pode prejudicar a utilidade do modelo.
- Passo 5 - Validação e Implantação: A validação do modelo é crucial para garantir seu desempenho adequado. A implantação sem validação adequada pode resultar em respostas incorretas ou enviesadas, prejudicando a

confiabilidade do sistema. A falta de melhoria contínua e ajustes finos pode levar a um modelo obsoleto.

Em resumo, a importância de seguir um procedimento bem definido e ajustável de implementação é fundamental para o sucesso da IA Generativa em qualquer organização. A não observância dessas etapas pode acarretar desafios e preocupações que variam de problemas legais a desempenho inadequado, prejudicando a eficácia da IA Generativa na solução de questões de conhecimento. Portanto, a adoção de boas práticas em todas as fases é crucial, e essas práticas podem ser adaptadas às necessidades específicas de cada organização, garantindo um resultado bem-sucedido.

Conclusão ou considerações finais

A implementação da Inteligência Artificial Generativa representa um marco significativo na jornada de transformação digital das organizações. Este artigo explorou os passos e os pontos-chave necessários para aproveitar plenamente o potencial dessa tecnologia revolucionária.

A IA generativa não é apenas uma ferramenta; é um catalisador da criatividade e da inovação. Com a capacidade de compreender e criar conteúdo em linguagem natural, essa tecnologia transcende as limitações tradicionais da IA e oferece oportunidades inigualáveis para melhorar processos, gerar *insights* e aprimorar a interação com clientes e parceiros.

Exploramos a importância da IA generativa como uma ferramenta que transcende a capacidade tradicional da IA, permitindo não apenas a compreensão, mas também a criação de conteúdo em linguagem natural. Abordamos os passos essenciais desse processo, incluindo autorização formal, definição de problemas, preparação de dados, treinamento de modelos e validação e implantação.

Em um ambiente empresarial em constante evolução, a IA generativa se destaca como uma ferramenta indispensável para superar desafios e alcançar um nível inédito de excelência. Ao adotar essa tecnologia de maneira estratégica, as organizações estarão bem posicionadas para prosperar, se destacar em suas respectivas

indústrias e atender às demandas cada vez mais complexas e diversificadas do mercado. A transformação digital com a IA generativa não é mais um cenário futuro; é uma realidade palpável e emocionante que aguarda aqueles que desejam liderar a inovação em suas organizações.

Referências e Sugestões Bibliográficas

- BERG, J.; RAJ, M.; SEAMANS, R. (2023) **Capturing value from artificial intelligence.** Academy of Management Discoveries, doi: 10.5465/amd.2023.0106.
- DE CREMER, D.; BIANZINO, N.M.; FALK, B. (2023) How generative AI could disrupt creative work. *Harvard Business Review*, 13 April, Disponível em: <https://hbr.org/2023/04/how-generative-ai-could-disrupt-creative-work>. Acesso em 27set2023.
- EDELMAN, D.C.; ABRAHAM, M. (2023) Generative AI will change your business: here's how to adapt. *Harvard Business Review*, 12 April. Disponível em <https://hbr.org/2023/04/generative-ai-will-change-your-business-heres-how-to-adapt>. Acesso em 27set2023.
- GILARDI, F.; ALIZADEH, M.; KUBLI, M. (2023) **ChatGPT outperforms crowd-workers for text-annotation tasks**, Working Paper, doi: 10.48550/arXiv.2303.15056.
- HWANG, H. G.; LEE, S. (2019). Artificial intelligence (AI) in business: a conceptual framework. *Futures* 109: 94-103
- JURAFSKY, D.; MARTIN, J. H. (2020) **Speech and Language Processing:** an Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Prentice Hall.
- KROLL, J. A.; KOHLI, N.; ROSA, H. (2019) Data protection by design: building the foundations of trustworthy data sharing. *International Data Privacy Law* 9(3): 167-184
-

Autor

Fábio Correa Xavier

CIO do TCESP. Professor e Coordenador Universitário. Colunista do MIT Technology Review Brasil e do IT Forum. Mestre em Ciência da Computação pela USP, MBA em Gestão de Negócios pelo IBMEC/RJ, e Especialização Network Engineering pela JICA-Japão, além de outras 4 pós-graduações. Certificações IAPP CIPM e CDPO/BR. EXIN Privacy and Data Protection. (ISC)² CC.

OS DESAFIOS DA INTELIGÊNCIA ARTIFICIAL NO SETOR PÚBLICO

Adriana Mecelis

Introdução

A revolução tecnológica está redefinindo muitos dos aspectos da sociedade e da vida do ser humano, de modo que a Inteligência Artificial (IA), juntamente com outras tecnologias emergentes, passa a ocupar um espaço de destaque na adequação de processos produtivos e de prestação de serviço, tanto para agentes públicos como privados. Mas a IA traz consigo riscos e incertezas que levam a grandes discussões, notadamente de caráter ético, em vista do seu grande poder de intervenção nas relações humanas.

Para o Estado, além do imenso desafio de incorporar o crescimento tecnológico para não deixar a prestação de serviço público à míngua de novas necessidades de uma sociedade digital, fica também o desafio de acompanhar o desenvolvimento desenfreado da tecnologia no setor privado, para agir em seu poder regulador quando necessário.

Neste artigo, examinaremos os desafios da IA no setor público, com um foco especial na produção de resultados de eficiência na prestação de serviço e na regulação estatal necessária para orientar e controlar seu uso.

Breve delimitação do conceito de IA

Existem múltiplas definições ou perspectivas para IA²¹. Para fins do presente artigo, utilizamos um conceito de ordem prática.

Primeiramente, destacamos que a IA aqui considerada se diferencia, pelo seu processo e seu resultado, de uma programação clássica. A confusão entre os conceitos é pertinente, uma vez que a programação clássica é ainda chamada de *IA baseada em regras*. Contudo, a *IA baseada em regras* constitui-se em um diálogo de *se – então (if – then)*²² em que não se identifica um processo de fato inteligente. É um processo que apresenta alto nível de explicabilidade, uma vez que o comando é dado pelo homem e não gerado pela máquina. Podemos dizer, então, que a partir de uma base de dados, há uma determinação humana, escrita em código computacional, para a realização automática de uma tarefa uma vez constatada como verdadeira uma dada premissa.

Esse tipo de automação é muito utilizado em processos produtivos e de facilitação de fluxos, inclusive na seara pública. No entanto, não apresenta profundos desafios de implantação e fiscalização, uma vez que o comando é dado pelo ser humano, ainda que em linguagem computacional, de modo que pode ser extraído e compreendido sem maiores dificuldades, possibilitando um controle nos moldes já realizados nos dias de hoje.

A IA como é concebida nas discussões atuais é diferente e pressupõe o conceito de aprender, que pode ocorrer por diferentes métodos de aprendizado de máquina (*machine learning*)²³. Assim, submetida a um grande conjunto de situações com aspectos em comum, a máquina investiga quais variáveis foram relevantes para a produção de um determinado resultado, criando o comando de

²¹ Para mais definições coletadas pela OCDE, consultar BERRYHILL, J. et al. Hello, World: Artificial Intelligence and its use in the Public Sector. OECD Working Papers on Public Governance Nº. 36. 2019. p. 13.

²² Na automação de um jogo de xadrez, por exemplo: • “SE a peça for um peão ENTÃO ela só poderá avançar uma posição em linha reta”

²³ Aprendizagem supervisionada, aprendizagem não supervisionada, aprendizagem por reforço, aprendizagem profunda, dentre outros.

execução para resultados futuros. Nesses casos, “os computadores não aprendem novos conhecimentos, eles aprendem como aprender” (OCDE, 2019, p. 47).

Isso consiste em dizer que esses novos dados passam a também a integrar a base de treinamento, transformando todo o processo de aprendizagem em um contínuo ciclo de retroalimentação, o que acarreta um aspecto de imprevisibilidade e incerteza de riscos, que deve ser cuidadosamente considerado quando da escolha da solução.

IA no setor público: o que está em jogo?

A constatação do proveito de qualquer ferramenta tecnológica deve sempre partir da pergunta sobre qual é o problema a ser resolvido e, a partir desta resposta, estudar as possíveis soluções que se apresentem, tecnológicas ou não, avaliando os benefícios e riscos envolvidos. Nos últimos anos, a IA tem sido a solução escolhida para muitos dos problemas que envolvem o setor público, visando a melhoria da eficiência, da tomada de decisões e da prestação de serviços aos cidadãos.

Um levantamento do Portal Transparência Brasil (2020) apontou a identificação de 44 ferramentas de IA usadas por órgãos governamentais, que foram agrupadas naquele estudo em duas dimensões:

- 1) ferramentas segundo sua utilização para tomada de decisão ou não;
- 2) ferramentas com público-alvo interno ou externo ao poder público (usuários finais).

A conclusão do levantamento pode ser visualizada na tabela a seguir:

CLASSIFICAÇÃO DE FERRAMENTAS DE INTELIGÊNCIA ARTIFICIAL (IA)
SEGUNDO USO PARA TOMADA DE DECISÃO E PÚBLICO-ALVO.

	COM TOMADA DE DECISÃO	SEM TOMADA DE DECISÃO	TOTAL
Público interno	20 ferramentas	16 ferramentas	36
Externo	8 ferramentas	0 ferramentas	8
Total	28	16	44

Fonte: Portal Transparência Brasil (2020)

Dentre as ferramentas identificadas, podemos destacar as utilizadas em diversas instituições para agilizar o atendimento do usuário de serviço público e para melhorar o fluxo processual e procedimental. É o caso dos *chatbots* e de ferramentas desenvolvidas para detecção de fraudes. Mas quais os riscos envolvidos nessas soluções?

Falar de riscos advindos da implementação da tecnologia pressupõe conhecer e considerar diferentes realidades culturais, financeiras, educacionais e geracionais, eventualmente existentes dentro de um mesmo contexto regional ou nacional.

Vejamos o exemplo do *chatbot*, ferramenta desenvolvida para auxiliar o atendimento de usuários de serviços públicos, otimizando o processo ao reduzir a necessidade de recursos humanos, reconhecidamente escassos no funcionalismo público. A sua implementação deve considerar o risco específico decorrente do “não acesso ao serviço por pessoas cuja inclusão digital vem sendo realizada de forma precária, incompleta ou inadequada, como analfabetas, imigrantes, deficientes ou sem acesso estável à internet” (PORTAL TRANSPARÊNCIA 2020, p. 11).

Portanto, a ferramenta tecnológica, quando concebida como um instrumento de agilidade para aqueles que têm amplo acesso à mecanismos digitais, não pode excluir outros usuários de seu direito à informação adequada e clara sobre o serviço pretendido.

Na detecção de fraudes, por sua vez, o risco é identificado a partir da possibilidade de conclusões falsas, quando embasadas em variáveis que não definem necessariamente uma relação causal com o resultado fraudulento. Esse é o contexto da chamada discriminação algorítmica, em que os dados históricos podem carregar questões discriminatórias dos mais diversos tipos, acabando por produzir um resultado injusto e eticamente desarrazoado ao impactar negativamente no acesso a serviços e exercícios de direitos.

Indo um pouco além neste exemplo, é emblemático o projeto CÉREBRO, desenvolvido pelo CADE para a detecção de carteis em licitações. Esse projeto foi concebido como um grande avanço ao permitir o cruzamento de dados de diversos processos licitatórios e, assim, identificar condutas reiteradas quanto à supressão de proposta, lances de cobertura, prevalência de desclassificação de lances, divisão de mercados, padronização de preços e similaridade nas propostas. Não obstante, também é possível identificar riscos em sua utilização, tais como o impacto a direitos de consumidores, investidores e concorrentes, na medida da possibilidade de distorção eventualmente ocasionada por dados não representativos da totalidade das situações em que será empregado.

Deste modo, a imprescindibilidade de avaliação adequada de riscos já se faz presente há um tempo, mas como ocorre em todo cenário que envolva novas tecnologias, a questão está em constante mudança e avanço. Quando, nos últimos meses de 2022 foi divulgado o acesso amplo (disponível a um público geral) ao CHATGPT, a IA Generativa se posicionou como principal alvo de estudo e desenvolvimento por parte de organizações públicas e privadas, trazendo novas perspectivas de uso e revelando incertezas e lacunas em regulações recém desenvolvidas sobre o tema.

Assim, os benefícios gerados pela utilização da IA se mostram exponencialmente crescentes, do mesmo modo que novas possibilidades de risco são verificadas, demandando a constante identificação de maneiras de mitigá-lo. A rapidez do desenvolvimento tecnológico, bem como o grande alvoroço causado pela abertura de ferramentas até então utilizadas em contextos controlados, desafiam cotidianamente o Poder Público, notadamente sob os aspectos que veremos a seguir.

Desafios operacionais, éticos e de privacidade na implantação e regulação estatal da IA

Os desafios para uma adequada utilização e regulação estatal da IA consideram questões comuns ao setor privado, mas que se revelam mais proeminentes a partir de peculiaridades existentes do setor público, conforme destacaremos a seguir.

Qualidade dos dados e proteção à privacidade

A implementação de processos baseados em IA pressupõe o tratamento massivo de dados, o que nos revela dois centros de preocupação de uso ético e eficaz da tecnologia:

1) a qualidade dos dados, ou seja, a utilização de bases de dados efetivamente representativas e apropriadas para o contexto de aplicação da ferramenta tecnológica; e

2) a privacidade dos dados e proteção dos cidadãos.

A questão é destacada no levantamento do Portal Transparência Brasil nos seguintes termos:

a estruturação dos bancos de dados de treinamento são questões centrais da análise, não só porque podem impactar direitos relacionados à privacidade e proteção de dados, mas também porque podem ser determinantes para formar ou não vieses potencialmente discriminatórios (TRANSPARÊNCIA, 2020. p.16).

Com a coleta e o armazenamento de dados pessoais em larga escala, os governos precisam adotar medidas rigorosas para proteger essas informações contra acessos não autorizados e violações de segurança. Além disso, conforme já explicitamos anteriormente, sistemas de IA que influenciam decisões governamentais podem ser suscetíveis a preconceitos, sendo essencial o treinamento dos algoritmos com dados imparciais, além de se garantir a existência de mecanismos de supervisão passíveis de identificar e corrigir qualquer viés que venha a surgir, assegurando aos cidadãos os seus direitos fundamentais.

Não se deve esquecer, por fim, que é da qualidade dos dados utilizados no treinamento que emergirá a qualidade de resposta dada pela IA e, assim, a efetividade da ferramenta, estando esse desafio diretamente conectado com a eficiência buscada na implementação tecnológica no setor público, seja para finalidade reguladora, seja para fins de prestação de serviço.

Estruturação Pública e capacitação dos funcionários públicos

Uma segunda questão que se coloca como o centro de atenção é capacitação e estrutura compatíveis com o desenvolvimento da tecnologia. Não é possível se falar em plena produção de resultados se não houver estrutura tecnológica ou humana adequada à ferramenta implementada.

A integração bem-sucedida da IA no setor público exige investimentos em infraestrutura de tecnologia da informação robusta e acesso a talentos qualificados no tema. Os governos precisam avaliar suas capacidades existentes e fazer investimentos estratégicos para preencher lacunas. Os funcionários públicos devem adquirir conhecimentos específicos para entender como a tecnologia funciona e como ela pode ser aplicada em suas áreas de atuação.

Assim, destinação de recursos e programas de treinamento e capacitação são fundamentais para garantir que a força de trabalho esteja preparada para a era da IA. Conforme destacado pela OCDE:

Mesmo quando a IA pode ser uma solução para problemas governamentais, a falta de técnicas básicas de gestão de dados pode limitar o seu potencial. Os funcionários públicos interessados em envolver-se com a IA precisam saber o que são dados, que tipos de dados podem ser utilizados, que tipo de dados a IA necessita e como verificar se os seus dados estão prontos para a IA (OCDE 2019, p.32 – tradução nossa)

Além disso, a colaboração com o setor privado é muitas vezes necessária para implementar a IA com sucesso no setor público, possibilitando que projetos já desenvolvidos e em utilização sejam aproveitados. Parcerias estratégicas podem fornecer acesso ágil a

recursos e tecnologias mais avançadas do que as desenvolvidas diretamente no contexto da Administração Pública.

Conscientização dos usuários

Outro ponto a ser verificado é a conscientização e familiaridade do destinatário do produto ou serviço desenvolvido com a IA, quanto aos aspectos que envolvam o resultado produzido pela ferramenta. Trata-se de questão intimamente ligada à noção de explicabilidade e transparência a ser abordada um pouco mais adiante. Contudo, aqui, trazemos a questão sob a perspectiva do usuário, já que não é possível se falar em controle social se não há capacidade de entendimento e averiguação da conduta tida por ilegal.

Trata-se de um grande desafio a ser operacionalizado pelo Poder Público, diante de diferenças regionais e sociais de acesso à informação, níveis de escolaridade e inclusão digital. Não obstante, é crucial o trabalho a ser desenvolvido, de modo a assegurar a autonomia de escolha daqueles que venham a ser atingidos, de forma direta ou indireta, pela ferramenta tecnológica.

Além disso, ações de conscientização auxiliam o Poder Público no controle de ações eventualmente danosas uma vez que, conscientes do dano sofrido ou presenciado, os usuários podem funcionar como agentes de comunicação, trazendo ao conhecimento do órgão regulador, por meio de requerimentos e denúncias, eventuais infrações que não sejam prontamente perceptíveis em um contexto mais amplo de fiscalização.

Transparência e explicabilidade

A transparência e explicabilidade da tecnologia pressupõe que os processos de tomada de decisão baseados em IA sejam justos e imparciais. A avaliação contínua do impacto da IA no setor público é essencial para garantir a constatação de sua eficácia e a conformidade com as leis e regulamentos existentes. Permite-se, com isso, o controle social sobre os atos praticados pelo Estado, conduzindo a questão em harmonia com os princípios que regem a atividade dos Poderes Públicos. São, neste aspecto, cruciais para determinar a

confiança nas instituições públicas, tornando possível a efetiva implementação de sistemas baseados em IA, pois “se o público não confiar que o governo utilize a IA de forma ética, poderá evitar serviços que utilizem IA ou opor-se à sua introdução.” (OCDE, 2019. p. 110).

Mas além do controle social, a transparência e a explicabilidade se destacam, inclusive no âmbito privado, na medida em que pressupõem um processo de desenvolvimento mais responsável, tendente a mitigar diferenças sociais e a opressão a grupos marginalizados. Ou seja, embora a exigência de sistemas transparentes possa esbarrar em questões de direito de propriedade intelectual, é dever do desenvolvedor não privar os destinatários do adequado esclarecimento acerca dos pontos prevalentes em uma tomada de decisão que lhe atinja direitos.

De modo que o processo de entendimento do funcionamento da IA e sua colocação em termos inteligíveis a terceiros (explicabilidade) também fomenta um desenvolvimento tecnológico socialmente sustentável, em que o ser humano é colocado como centro das ações adotadas.

Ser humano no centro das decisões acerca da implementação de IA

Por fim a colocação do ser humano como centro da preocupação das ações governamentais que visam a implantação de ferramentas de IA encontra coerência com os objetivos de políticas públicas em seus diversos campos de implementação. O conhecimento das necessidades daqueles que serão atingidos, direta ou indiretamente é ponto crucial para a implementação da IA, sendo o usuário o centro das ações governamentais. No Brasil, a Estratégia de Governo Digital para o período de 2020 a 2023, instituída pelo Decreto nº 10.332/2020 e pelo Decreto nº 11260/2022 define o Governo Inteligente como aquele que:

implementa políticas efetivas com base em dados e evidências e antecipa e soluciona de forma proativa as necessidades do cidadão e das organizações, além de promover um ambiente de negócios competitivo e atrativo a investimentos²⁴.

Portanto, a revolução tecnológica, principalmente no que diz respeito a ações governamentais, deve se pautar na melhoria da vida das pessoas, sem as colocar em um papel secundário, em que as ferramentas tecnológicas assumam o protagonismo em detrimento das necessidades individuais ou coletivas inerentes aos seres humanos.

Conclusão

Este artigo explorou as implicações da IA nas atividades do setor público, trazendo o conceito desta tecnologia, questões atreladas à identificação de riscos e desafios operacionais, éticos e de privacidade associados à sua implementação. A análise apresentada evidencia a necessidade de uma abordagem cuidadosa e informada por parte dos governos já que a IA, ferramenta poderosa e em constante desenvolvimento, traz, juntamente com benefícios significativos, riscos consideráveis, muitas vezes sequer ainda conhecidos.

A qualidade dos dados, a privacidade dos cidadãos, a necessidade de conscientização dos destinatários, a estruturação física e a capacitação dos funcionários públicos, bem como a transparência e explicabilidade da IA emergem como desafios críticos que devem ser considerados para assegurar a confiança no sistema e gerar resultados qualificados e imparciais. Neste ponto, colocar o ser humano no centro das decisões mostra-se crucial para garantir que as tecnologias

²⁴ BRASIL. Decreto nº 10.332, de 28 de abril de 2020. Institui a Estratégia de Governo Digital para o período de 2020 a 2022, no âmbito dos órgãos e das entidades da administração pública federal direta, autárquica e fundacional e dá outras providências. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2020/decreto/d10332.htm. Acesso em: 24 jul. 2022

sirvam para melhoria da qualidade de vida das pessoas e não as releguem a um papel coadjuvante.

Assim, em um mundo em que a IA se torna cada vez mais onipresente, é essencial que os governos atuem de forma proativa não só para implantar a tecnologia em seus serviços digitais, mas também para regulamentá-la, fiscalizando e orientando seu uso. A IA tem o potencial de revolucionar positivamente a prestação de serviços públicos e privados, mas apenas se for gerenciada com responsabilidade e sensibilidade às necessidades e aos direitos fundamentais do ser humano.

Referências Bibliográficas

BERRYHILL, J. et al. (2019) **Hello, World: Artificial Intelligence and its use in the Public Sector**. OECD Working Papers on Public Governance 36. Disponível em https://www.oecd-ilibrary.org/governance/hello-world_726fd39d-en. Acesso em 04 de out de 2023.

ENAP - Escola Nacional de Administração Pública (2019) **Prevenção e detecção de Cartéis em Licitações Módulo 3. Repressão a Cartéis**.

Disponível em: https://repositorio.enap.gov.br/jspui/bitstream/1/6495/3/Mod%203_Repress%C3%A3o%20a%20Cart%C3%A9is.pdf . Acesso em 04 de out de 2023.

OECD. **Artificial Intelligence in Society**. Recommendation of the Council on Artificial Intelligence. Disponível em: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>. Acesso em 02.10.2023.

TRANSPARÊNCIA BRASIL (2020) **Recomendações de Governança: Uso de Inteligência Artificial pelo Poder Público**. Disponível em: https://www.transparencia.org.br/downloads/publicacoes/Recomendacoes_Governanca_Uso_IA_PoderPublico.pdf. Acesso em 23 set. 2023.

Autora

Adriana Mecelis

Mestra em Direito e Tecnologia pela Escola de Direito da Fundação Getúlio Vargas. Especialista em Direito Público e em Direitos Difusos e Coletivos. Procuradora Federal.

MINERAÇÃO DE PROCESSOS E A ADMINISTRAÇÃO PÚBLICA

Sarajane Marques Peres
Marcelo Fantinato
Renata Mendes de Araujo
Fernando Henrique da Silva Costa

Introdução

A mineração de processos promove uma sinergia entre ciência de dados e ciência de processos. Essa sinergia facilita o desenvolvimento de ferramentas teóricas e práticas que melhoram a precisão, confiabilidade e eficácia da gestão de processos de negócio. Para contextualizar essa gestão, considere que um processo de negócio consiste em um conjunto de atividades que são executadas de forma coordenada, dentro de um ambiente organizacional e técnico (WESKE 2012). Essas atividades são executadas com o objetivo de alcançar metas de negócio. Processos de negócio, nesse contexto, abrangem diferentes domínios de aplicação (administração, operações de manufatura, monitoramento de jornadas etc.) e tipos de organização (pública ou privada, de natureza econômica ou social, educacionais, religiosas etc.).

Já a gestão de processos de negócio inclui conceitos, métodos e técnicas que apoiam o projeto e modelagem, a administração, a configuração, a melhoria e a análise de processos de negócio (WESKE 2012). A mineração de processos nasceu dentro da área de gestão de processos de negócio a partir da implementação de estratégias que permitem usar dados sobre processos e modelos de processo de maneira cooperativa, em análises focadas principalmente na noção de instâncias de processo (execuções de um processo), ou,

como é comumente denominado em mineração de processos, os “casos”.

Para exemplificar como um processo de negócio é tratado no âmbito da mineração de processos, considere o exemplo da figura 1. Na figura, um processo de negócio referente à gestão de concessões é apresentado, sob a representação lógica de uma notação chamada BPMN (do inglês *business process model notation*), constituindo-se em um modelo de processo (hipotético) para solicitações de concessão. A leitura deste modelo permite entender a lógica de trabalho usada para tratamento de solicitações de concessão. Por exemplo, a seguinte sequência de atividades pode ocorrer dada essa modelagem: *receber solicitação de concessão, coletar informações do requerente, analisar informações, coletar documentos comprobatórios, analisar documentos, verificar requisitos, negar concessão*.

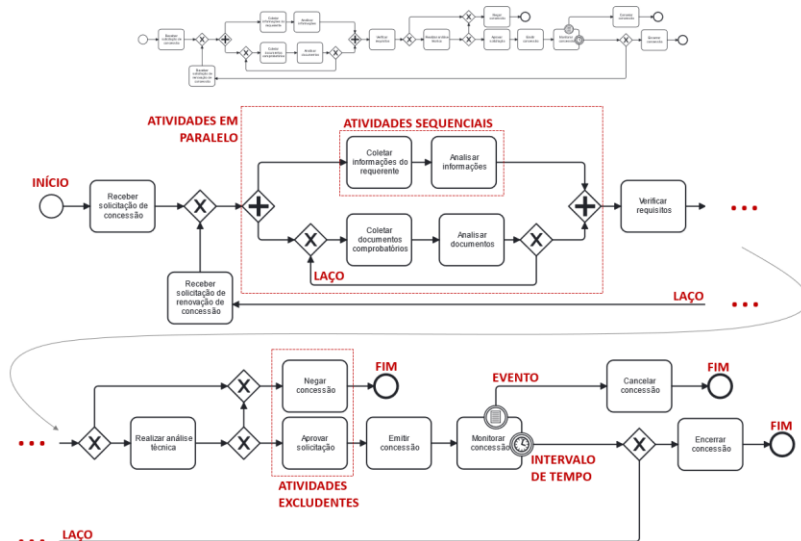


Figura 2: Exemplo de um modelo de processo na notação BPMN: processo de solicitação de concessão. O modelo completo é apresentado na parte superior da figura. Recortes do modelo são apresentados em destaque para permitir a visualização adequada. Anotações em vermelho trazem informação adicional (não fazem parte da sintaxe da notação).

O ponto crucial no contexto da mineração de processos é que, infelizmente, um modelo de processo de negócio bem estruturado e claro como o modelo da figura 1 nem sempre está disponível para quem gerencia o processo. Ainda, quando disponível, duas questões são importantes: a primeira é que a realidade de uma organização não está, necessariamente, refletida corretamente em um modelo como este; a segunda é que, sem ferramentas automatizadas e com algum nível de sofisticação, os modelos de processo têm sua utilidade limitada. O objetivo da mineração de processos é aumentar a compreensão de um determinado processo de negócio por meio da análise automática de dados sobre a dinâmica de um processo de negócio, seja descobrindo um modelo de processo aderente à realidade do trabalho executado na organização, seja usando esses dados e os modelos descobertos ou existentes de forma eficiente e eficaz para otimizar a gestão organizacional.

Processos permeiam todo o funcionamento da administração pública, sejam eles processos operacionais internos às agências, de prestação de serviços ao cidadão, de tomada de decisão ou de formulação de políticas públicas. Processos têm sido objeto de interesse das iniciativas de transformação digital no setor público, exigindo avanços tecnológicos, investimento em capacitação, definição clara de estratégias e mudanças significativas das visões de gestão pública (JANOWSKI 2015, FILGUEIRAS *et. al.* 2019, LEI 12.527 2011, LEI 14.129 2021). A transformação digital na administração pública não acontece com a simples introdução de novas tecnologias em seus processos, mas com abordagens que permitam uma revisão, reflexão e mudança dos processos e de seus colaboradores.

A gestão de processos de negócio surge como abordagem importante para o alinhamento da administração pública a esses desafios, tanto para gestão e melhoria de processos, quanto para auditoria e transparência, facilitando o acesso à informação sobre os processos (tramitação, regras, documentos, desempenho etc.) e promovendo a transparência e participação cidadã em ações de melhoria e inovação. Como princípios fundamentais da administração pública, a execução de processos requer o cumprimento de exigências de legalidade, impessoalidade, moralidade, publicidade e eficiência

(LIMPE). Mais recentemente, nessa mesma linha, as agências públicas têm sido desafiadas pela necessidade do *compliance*, também referido como conformidade legal. Em especial, o princípio da eficiência e o *compliance* podem ser bastante beneficiados pela mineração de processos.

Este texto pretende apontar para as oportunidades da mineração de processos como ferramenta para apoiar a administração pública no cumprimento de seus valores e princípios, bem como atender às exigências atuais de digitalização, de novos modelos de formulação de políticas públicas baseadas em dados, de desempenho, *compliance* e transparência de seus processos. Para isso, o texto é organizado a partir da seguinte sequência de tópicos: log de eventos, a fonte de informação primária para a mineração de processos; tarefas da mineração de processos e seus propósitos, exemplificados no processo (hipotético) de solicitação de concessão apresentado nesta introdução; aspectos cruciais da responsabilidade em mineração de processos; e iniciativas de aplicação de mineração de processos no contexto internacional e nacional da administração pública.

Fonte de informação: *log de eventos*

O principal insumo para a mineração de processos está relacionado a dados extraídos da dinâmica de execução de um processo de negócio, enquanto ela ocorre. Esses dados são obtidos a partir dos sistemas de informação que apoiam os processos de negócio dentro de uma organização.

Um *log de eventos* padrão é um arquivo sequencial no qual são registrados os dados sobre os eventos relacionados à execução de atividades referentes a um processo de negócio. Normalmente, assume-se que cada evento está relacionado a um caso específico (ou seja, uma instância de processo) e possui os atributos básicos identificador, *timestamp* (momento em que o evento ocorreu no formato de data e hora), atividade associada, assim como atributos adicionais não obrigatórios como custo e recurso (de *software* ou humano) associados à execução do evento. O caso, por sua vez, também possui atributos que descrevem o contexto no qual ele ocorre, incluindo um atributo obrigatório chamado *traço* (do inglês

trace), que especifica a sequência de eventos que compõem o caso. A individualidade dos traços é conhecida como variante de um processo. A figura 2 ilustra um trecho de um log de eventos com eventos agrupados por caso e ordenados pelo *timestamp*.

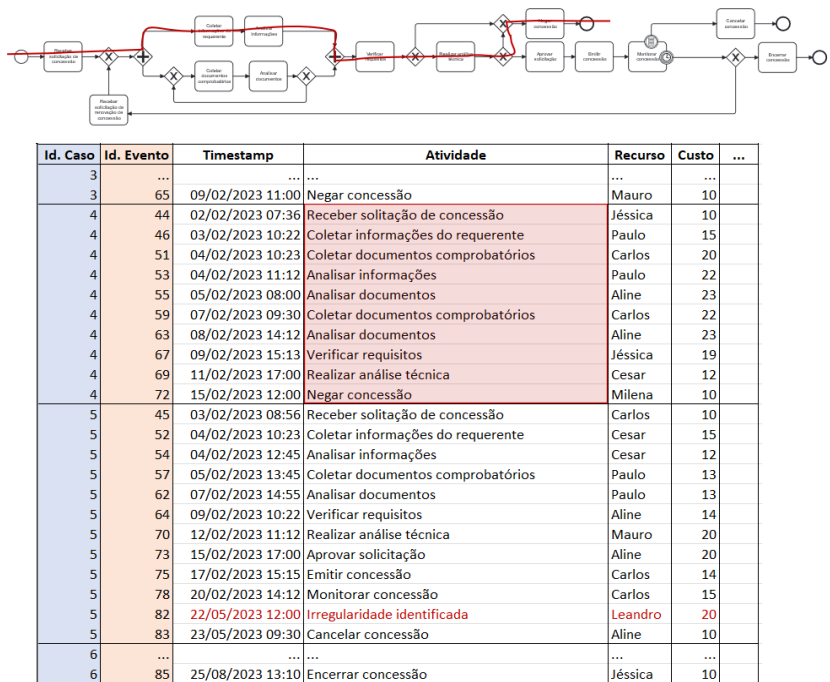


Figura 3: Trecho de um log de eventos correspondente ao modelo de processo apresentado na figura 1, o qual é destacado novamente na parte superior desta figura. Cada evento no log corresponde à execução de uma das atividades do modelo de processo, conforme o nome constante na coluna “Atividade” do log de eventos. O evento 82, no caso 5, diferentemente dos demais eventos do log, não se refere a uma atividade do modelo de processo, mas sim a um evento do modelo de processo de acordo com a notação BPMN. O caminho destacado em vermelho no modelo de processo na parte superior da figura corresponde ao caso 4 destacado também em vermelho no log de eventos. Observe que os identificadores de eventos são únicos dentro do log de eventos e associados conforme o momento de execução daquele evento.

A necessidade de estruturar um log de eventos sobre um único processo de negócio limita a representação de situações do mundo real. A fim de possibilitar a evolução da mineração de processos no que diz respeito à amplificação da visão dos processos de uma organização, foi estabelecido um outro padrão de log de eventos

chamado de log de eventos centrado em objetos (*object-centric event log* – OCEL²⁵). Esse padrão de *log* de eventos se concentra no registro de eventos associados a atividades que, por sua vez, estão associadas a diferentes objetos tratados dentro dos processos de uma organização de uma forma geral. Como efeito, um *log* de eventos do tipo OCEL diz respeito a mais do que um processo de negócio, permitindo a análise de vários processos e suas relações.

Organização da mineração de processos: tarefas e propósitos

Conforme declarado por Van Der Aaslt (2022), as perspectivas de análise de um processo de negócio via mineração de processos podem ser “direcionadas ao passado” (*backward-looking*) ou “direcionadas para o futuro” (*forward-looking*). Análises retrospectivas facilitam a derivação de modelos descritivos que revelam regras e comportamentos empregados no passado. Por outro lado, análises voltadas para o futuro potencializam a construção de modelos preditivos e prescritivos, possibilitando a otimização dos processos. Para a implementação dessas análises, a disciplina de mineração de processos foi organizada originalmente em três grandes tipos de tarefas (VAN DER AASLT, 2016): descoberta de modelo de processo, verificação de conformidade e melhoria de processo. Recentemente, Van Der Aaslt (2022) propôs uma organização mais granular cujas tarefas são:

- Descoberta de modelo de processo: engloba a proposição de algoritmos capazes de analisar dados históricos sobre um processo, registrados na forma de *logs* de eventos, e derivar um modelo lógico descritivo daquele processo, em uma linguagem de especificação apropriada. Entende-se um modelo de processo descritivo como aquele que explicita o que realmente ocorre durante a execução de um processo, ou seja, descreve o processo como “ele é” e não como ele “deveria ser” (o modelo de processo como “deveria ser” é dito normativo e, em geral, é desenhado conforme as necessidades organizacionais ou conforme regras e

²⁵ <https://ocel-standard.org>

leis associadas). A figura 3 mostra um cenário com um modelo de processo descritivo e um modelo de processo normativo, ambos hipotéticos. Essa tarefa de descoberta de modelo de processo é útil para tornar conhecida a real lógica da prática de trabalho em execução em uma organização e permitir uma avaliação detalhada dessa prática. Isso permite que gestores e outras partes interessadas estejam claramente cientes do processo em execução e habilitados para o gerenciar efetivamente, promovendo transparência sobre o trabalho realizado na organização.

- Verificação de conformidade: em essência, trata da comparação entre os dados de um log de eventos e a lógica de execução de um processo dada por um modelo de processo normativo. Disparidades entre o que está registrado no log de eventos e o que era esperado diante do modelo de processo normativo indicam ou que ações que violam regras e normas foram realizadas no passado, ou que há práticas de trabalho sendo executadas que não estão previstas no modelo, o qual precisa ser atualizado. A atividade de comparação de um log de eventos contra um modelo de processo pode assumir diferentes papéis dentro da prática de mineração de processos. Do ponto de vista dos projetistas de algoritmos para descoberta de modelo de processo, a comparação do modelo de processo descritivo descoberto contra um log de eventos usados para teste permite medir, sob algumas perspectivas, a qualidade do algoritmo criado. Do ponto de vista da gestão de processos, a comparação do modelo de processo normativo contra um *log* de eventos revela comportamentos reais destoantes do esperado e pode ser usada, por exemplo, para implementar procedimentos para *compliance*, auditoria e alinhamento de negócios, podendo inclusive facilitar o atendimento a requisitos de legalidade. O cenário ilustrado na figura 3 associa, de forma didática, a execução de uma ação de verificação de conformidade e de *compliance*.

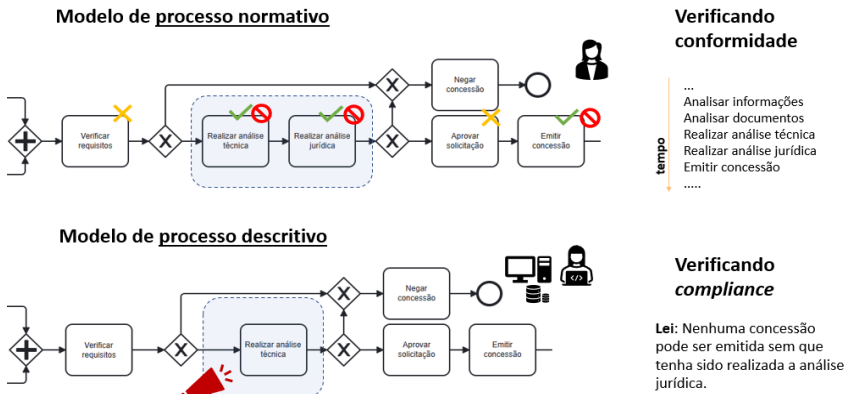


Figura 4: Um trecho de modelo de processo normativo (acima) e um de trecho de modelo de processo descritivo (abaixo) são apresentados em notação BPMN. O modelo descritivo descoberto a partir de um log de eventos informa que as instâncias de processos já executadas não realizaram a atividade “análise jurídica” durante o processo de solicitação concessão. Em uma verificação de conformidade ilustrativa que compara o modelo normativo com o trecho de log de eventos mostrado no canto superior direito da figura, o símbolo ✕ no modelo de processo indica que algumas atividades esperadas não foram executadas, enquanto os símbolos ✓🚫 juntos indicam que algumas atividades foram executadas sem que atividades anteriores entendidas como pré-requisitos tivessem sido executadas. Considerando uma verificação de compliance, o símbolo 🚨 no modelo descritivo representa um alerta de non-compliance, uma vez que a atividade “Realizar análise jurídica” não foi executada como exigido por lei (canto inferior direito). Quando o modelo descritivo foi contrastado com a lei em uma ação de auditoria, observou-se que no passado concessões podem ter sido emitidas sem a execução da atividade de análise jurídica, a qual, na realidade, era esperada em sequência à atividade de análise técnica de acordo com o modelo normativo.

- **Análise de desempenho:** está relacionada à análise de dados temporais, para que gestores possam verificar a eficiência com a qual o processo de negócio vem sendo executado. Usualmente, nessa análise, são criados painéis de indicadores-chave de desempenho (do inglês KPI – *key performance indicators*) e são revelados os ditos “gargalos” do processo, que são situações nas quais há uma ineficiência no trabalho realizado por fatores como acúmulo de tarefas, falta de infraestrutura, problemas logísticos etc.. Na figura 4, as setas mais grossas (em vermelho) indicam

uma lentidão no processo, identificada por uma análise de tempo realizada sobre os *timestamps* associados aos eventos no *log* de eventos. A informação trazida pela análise de desempenho pode apoiar também o estabelecimento de formas de solucionar os gargalos, uma vez que torna clara a situação de ineficiência de forma associada à lógica de trabalho explícita em um modelo de processo descritivo.

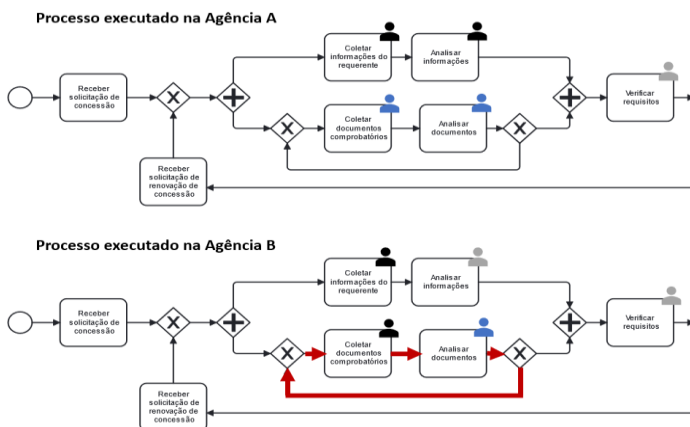


Figura 5: Modelos de processo descritivos nos quais a espessura das flechas indica o tempo entre o início da execução de cada atividade (quanto mais grossa a flecha, maior o tempo). Os modelos estão anotados com a alocação de recursos humanos que executam determinadas tarefas. Na Agência A, a coleta de informações do requerente e respectiva análise é realizada pela mesma pessoa, enquanto outra pessoa atua na coleta e análise de documentos comprobatórios. Na Agência B, uma mesma pessoa coleta informações do requerente e documentos comprobatórios, enquanto outras duas pessoas diferentes realizam as duas respectivas atividades de análise. De acordo com o exemplo, o processo na agência A é mais eficiente (observe que há um laço de execução no modelo de processo da Agência B que possui as flechas mais grossas, destacado em vermelho). A forma de distribuição dos recursos pode ser investigada para confirmar uma hipótese de que ela é a

- Mineração de processos comparativa: promove análises que contrastam dados e comportamentos de processos de negócio similares, e promove a transferência de conhecimento entre organizações, filiais ou departamentos. Trata-se de uma tarefa conhecida como *benchmarking* em prol da otimização dos processos. A forma de implementação dessa tarefa pode variar de acordo com as necessidades e possibilidades de cada cenário. A

comparação pode ocorrer entre dois ou mais modelos de processo descritivos (Figura 4), ou na forma de verificação de conformidade, contrastando *logs* de eventos contra modelos descritivos ou normativos provenientes de diferentes lugares.

Rotinas repetitivas

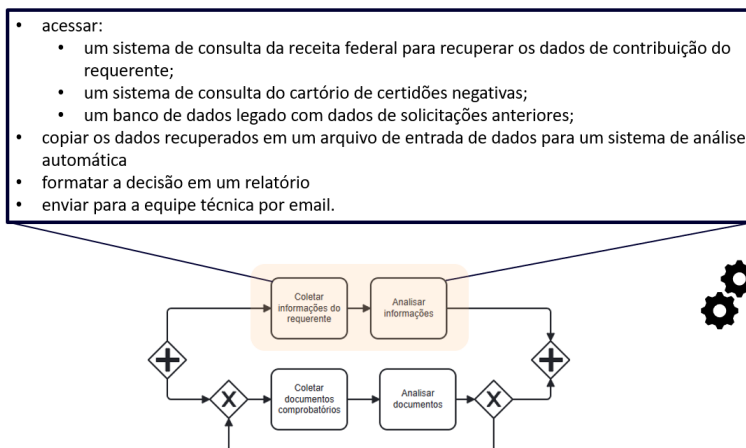


Figura 6: A mineração de tarefas (task mining) consiste em analisar cada tarefa realizada para executar atividades de um processo de negócio. Exige dados mais granulares do que aqueles presentes em um log de eventos tradicional. Na figura, o quadro superior destaca um conjunto de tarefas realizadas enquanto as atividades “Coletar informações do requerente” e “Analisar informações” são executadas. Várias dessas tarefas podem ser automatizadas por serem consideradas “robóticas”, de forma a liberar tempo de trabalho de um recurso humano.

- Mineração de processos orientada a ações: promove a automação da gestão de processos (apoando mineração de tarefas e automação de processos robóticos), provendo informação para a implementação da transferência de responsabilidade de determinadas sequências de ações de um recurso humano para um recurso computacional (veja um exemplo na figura 5). Em geral, ações rotineiras como acessar uma base de dados, realizar consultas em *websites*, copiar dados e enviar para um sistema, se realizadas com muita frequência, consomem um tempo precioso no dia de trabalho de um colaborador e, ainda, por serem simples e enfadonhas, não

recebem a devida atenção gerando um risco de produção de erros. Um ponto interessante dessa tarefa é que resultados da execução de outras tarefas de mineração de dados, como a descoberta de modelo de processo e a verificação de conformidade, podem gerar informações preciosas que permitem identificar, com precisão, sequências de tarefas que podem ser automatizadas.

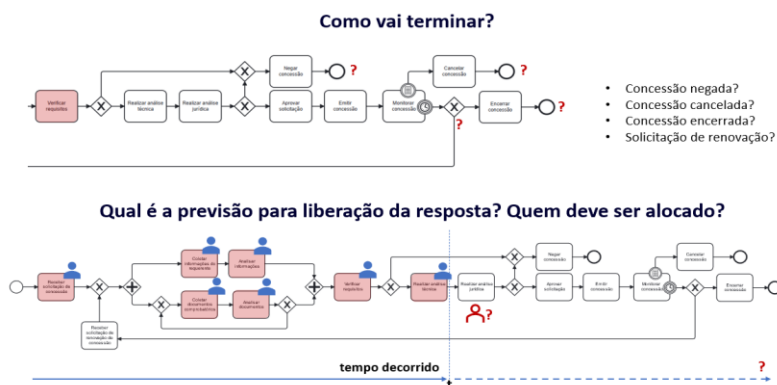


Figura 7. Na parte superior, o retângulo colorido, referente à atividade “verificar requisitos”, indica que uma instância de processo sob monitoramento já executou essa atividade. Neste momento, quatro desfechos são possíveis. Indicar qual deles é o mais provável, com base em um histórico de instâncias de processos já finalizadas caracteriza a tarefa de predição de desfecho de um caso. Na parte inferior, várias atividades foram executadas para uma determinada instância de processo, como indicado pelos retângulos coloridos. A atividade que acabou de ser executada é a atividades “realizar análise técnica”. Nesse cenário, duas tarefas estão ilustradas na figura: a predição de quanto tempo ainda demorará para que essa instância de processo seja finalizada; e a tarefa prescritiva de recomendar um recurso humano adequado para executar a próxima atividade (“realizar análise jurídica”) do processo. Nesse último caso, o objetivo é realizar uma recomendação que otimize a execução daquela instância de processo em determinado aspecto (eficiência, por exemplo), dadas determinadas restrições (carga atual de trabalho dos recursos humanos e legislação do trabalho, por exemplo).

- Mineração de processos preditiva: implementa modelos preditores para adiantamento de fatos, comportamentos e tendências. Essa tarefa é principalmente importante quando há a necessidade de realizar monitoramento de processos, já que ela

permite implementar “alarmes” para evitar ocorrências indesejadas ou para indicar a necessidade de ações corretivas, e pode apoiar decisões que mudem um curso de ação, se adequado. Exemplos de predições que são comumente estudadas na mineração de processos são: predição de tempo de resolução de um caso em que, a partir da execução ainda parcial de uma instância de processo, é possível estimar quanto tempo ainda falta para que ele seja concluída; predição da próxima tarefa em que, também em meio à execução de uma instância de processo, um modelo preditor pode dizer qual tarefa, dentro de um conjunto de possibilidades, é a mais provável de ser executada naquele momento; e ainda, predições referentes a desfechos prováveis na execução de um caso são antecipados, permitindo, por exemplo, preparar-se, adequadamente, para outras atividades subsequentes necessárias naquele contexto de desfecho. A figura 6 ilustra alguns cenários de predição.

Além dessa classificação, a mineração de processos tem apresentado demandas para tarefas prescritivas (KUBRAK *et al.* 2021), como recomendação de atividades e de recursos (Figura 6). A recomendação de atividades é pertinente a ambientes nos quais os processos são altamente dependentes de conhecimento, possivelmente conhecimento tácito. A recomendação de recursos diz respeito à implementação de sistemas que indicam quem, ou qual máquina ou ferramenta de *software*, deve ser alocado para realizar uma tarefa, considerando um conjunto de restrições existentes no ambiente e no contexto do processo de negócio.

Mineração de processos responsável

Mineração de processos é uma disciplina que tem dados como entrada e conhecimento como saída. Em particular, os dados que fomentam as análises de mineração de processos dizem respeito a dinâmicas de trabalho, estão associados a recursos humanos e sistêmicos envolvidos nesta dinâmica e carregam consigo informações que, direta ou indiretamente, dizem respeito a individualidades, inclusive comportamentais. Além disso, depois de processados pelos

diferentes algoritmos e métodos de mineração de processos, esses dados se transformam em conhecimento de valor agregado que apoiam decisões cujos efeitos influenciam pessoas e organizações e trazem algum tipo de destaque, ou poder, àqueles que detêm esse conhecimento e o privilégio de usá-lo. Nesse contexto, os mais diversos riscos éticos, morais e técnicos emergem, e alguns princípios devem nortear o desenvolvimento e aplicação dos recursos dessa área. A figura 7 organiza alguns dos princípios e ações que são imprescindíveis no contexto da mineração de processos.

Imparcialidade	Prevenir-se de conclusões injustas e <i>non-compliant</i> , mesmo se apoiada por dados e algoritmos
	Comunicar e estar ciente dos vieses presentes nos dados e em políticas de algoritmos
Acurácia	Fornecer medidas interpretáveis para precisão
	Comunicar e estar ciente da limitação de medidas, métodos e artefatos
Confidencialidade	Respeitar a privacidade e o direito dos clientes, funcionários e outros participantes do processo
	Proteger dados confidenciais pessoais e não pessoais
Transparência	Fornecer rastreabilidade total à fonte de dados, algoritmos e medidas
	Comunicar e estar ciente dos problemas de qualidade de dados e limitação dos algoritmos e medidas

Figura 8: Princípios e ações da mineração de processos responsável. Resumo baseado nas discussões apresentadas por VAN DER AALST *et al.* (2017).

A imparcialidade, também conhecida como “justiça”, está relacionada à noção de vieses e preconceitos. A ideia é que seja possível evitar conclusões e ações injustas, ou que não estejam alinhadas com regras e leis, ainda que tais ações tenham sido derivadas de forma “precisa” por algoritmos. Este é um ponto delicado para quem lida com áreas em que os algoritmos se baseiam em dados históricos. Os algoritmos, em sua maioria, modelarão o comportamento presente nos dados e, portanto, reproduzirão vieses e conceitos neles embutidos pelo histórico organizacional. Além disso, otimizações buscadas por esses algoritmos estarão focadas nas restrições que lhes forem impostas. Toda e qualquer conclusão ou ação que venha a ter efeito no ambiente de uma organização precisa ser verificada antes de ser divulgada ou colocada em prática.

A acurácia, enquanto princípio, precisa ser tecnicamente mensurável, compreendida e interpretável. Obviamente que os resultados oferecidos pela mineração de processos não são puramente suposições. Estatísticas apropriadas estão atreladas aos resultados. Contudo, medidas e estatísticas são facilmente mal

interpretadas e frequentemente não interpretáveis no sentido contextual de um processo de negócio. Além disso, na mineração de processos há duas questões que precisam ser cuidadosamente tratadas: o conhecimento acurado sobre as linguagens de representação, uma vez que elas possuem limitações de expressividade que podem levar a entendimentos errôneos sobre a dinâmica de trabalho; e as limitações e vieses dos algoritmos que, em geral, possuem regras heurísticas associadas e métodos de otimização aproximados. A mineração de processos, como uma disciplina apoiada no raciocínio indutivo, pode incorrer em falhas por princípio, uma vez que um argumento indutivo “é aquele no qual é pretendido que seja improvável, porém possível, que existam premissas verdadeiras levando a conclusões falsas” (COOK 2009, 150).

Confiabilidade está ligada à privacidade e segurança. As análises propiciadas pelas informações presentes no *log* de eventos, sejam elas análises descritivas, preditivas ou prescritivas, podem revelar, além da dinâmica de execução das atividades, o comportamento de pessoas associadas à essa dinâmica. Uma análise referente ao desempenho de um processo de negócio revela características de desempenho associadas aos recursos humanos que executam as atividades daquele processo e podem, eventualmente, serem usadas como forma de avaliação indireta e repreensão, sem que a “vigilância” fosse de conhecimento daqueles recursos humanos. Ainda, uma eventual falta de cuidado com a segurança de um *log* de eventos ou dos algoritmos que os processam pode revelar informações sensíveis associadas a pessoas ou à sua relação com determinadas atividades.

A transparência é o princípio que transmite a sensação de “honestidade” de uma abordagem. Nesse sentido, espera-se que os métodos usados para produção dos resultados sejam reproduzíveis e que os resultados dos exercícios de reprodução sejam significativamente coerentes com os resultados originalmente apresentados. Complementarmente, sempre que existe a possibilidade da obtenção de resultados diferentes, ainda que em ações de reprodutibilidade, é necessário que essa condição de insegurança seja medida e clara e detalhadamente reportada a quem é de interesse.

Iniciativas pelo mundo e no Brasil

A mineração de processos aplicada na administração pública ainda está restrita a poucas iniciativas. Muito embora a utilidade dessa combinação possa ser bem compreendida e aceita, a inserção dessa prática como recurso de apoio administrativo requer investimento financeiro para formação de recursos humanos habilitados à realização da análise e instrumentação de sistemas de informação que possibilite a construção de *logs* de eventos e, em alguns casos, uma tramitação legal de autorização de realização da análise, já que muitos processos públicos envolvem dados e informações sensíveis referentes aos cidadãos. Uma pesquisa na base de indexação de artigos científicos Scopus²⁶ permite a criação de um quadro ilustrativo de como a administração pública, em diferentes países, tem usado mineração de processos. Descoberta de modelo de processo (BASANYA 2011, STEFANINI 2016, REPTA 2018) e verificação de conformidade (PETRUSEL 2013, KALENKOVA 2018, ZAMAN 2020, SANGIL 2020, GONZÁLEZ 2021, GONZÁLEZ 2021) são as tarefas que receberam mais atenção no contexto da administração pública, mas análise comparativa (LARSSON 2021) e construção de modelo de predição (WESTPHAL 2020) também estão presentes neste quadro, considerando os trabalhos realizados fora do Brasil.

Na iniciativa reportada por Basanya (2011), a descoberta de modelo de processo é implementada dentro de uma solução cujo objetivo é descobrir, de forma automática, quais serviços e processos oferecidos nas plataformas do governo possuem relações que os caracterizem como candidatos a comporem uma solução geral e colaborativa de administração pública. Interessantemente, essa iniciativa não possui um *log* de eventos à disposição e usa mineração de textos para fomentar os algoritmos que descobrirão as relações dentro da dinâmica dos processos. Stefanini (2016) e Repta (2018) seguem uma linha mais tradicional. A primeira iniciativa propôs um método para descoberta da lógica presente no fluxo de pacientes e relata que a compreensão dessa dinâmica permitiu modelar de forma

²⁶ <https://www.scopus.com/>

adequada o sistema de apoio à gestão de saúde. A segunda iniciativa usou ferramentas de mineração de processos já integradas aos sistemas do governo para descobrir como ocorre o fluxo de documentos nos trabalhos administrativos.

Verificação de conformidade foi aplicada por: Petrusel (2013), com uma ferramenta para compreender e melhorar os processos de emissão de licenças de construção; Kalenkova (2018), para compreensão das divergências existentes entre o comportamento real do usuário e o comportamento esperado a partir do ponto de vista dos desenvolvedores dos sistemas que apoiam os serviços de governo eletrônico (registros de carros, solicitação de passaportes, consultas médicas e pagamentos de multas e impostos); Zaman (2020), em uma iniciativa de análise de adesão à disposições da GDPR (do inglês *General Data Protection Regulation*), com especial atenção às cláusulas relacionadas aos direitos dos titulares dos dados; Sangil (2020), na promoção da transparência pública, explorando a verificação de conformidade para descobrir como alguns processos públicos nas Filipinas foram executados e verificar sua conformidade com as leis relacionadas, tendo relatado a ampliação da transparência, eficiência e responsabilidade geral, permitindo identificar comportamentos inadequados, por menores que fossem os desvios; e González *et al.* (2021) e González *et al.* (2022), em uma análise de processos relacionados à emissão de passaporte por meio de verificação de conformidade para encontrar violações de *compliance* em um sistema de governo eletrônico implementado em um ambiente estendido de processos de negócio colaborativos.

Uma iniciativa que constrói um modelo preditivo foi apresentada por Westphal (2020). O tempo de espera por atendimento em um departamento de emergência foi desenvolvido a partir de um modelo de aprendizado de máquina treinado sobre informações de um *log* de eventos que continha a jornada do paciente, incluindo dados sobre ele e seu atendimento. Por fim, a análise comparativa entre processos de negócio foi aplicada por Larsson (2021) para analisar a dinâmica de atendimento de usuários via sistemas de associação automática a benefícios socioeconômicos e de usuários que se inscrevem de maneira manual para recebimento dos benefícios.

No Brasil, foram identificadas três iniciativas relatadas na literatura acadêmica. Costa e Rodrigues (2020) aplicaram mineração de processos para analisar processos da Marinha do Brasil, mostrando o potencial para descoberta da execução real de processos, gargalos e desvios. Os autores reportaram que algumas ineficiências identificadas podem ser resolvidas com ações pontuais de reorganização de trabalho. Unger *et al.* (2021) aplicaram mineração de processos descritiva, incluindo a descoberta de modelo de processo, em um *log* de eventos referente a movimentos de processos jurídicos do Tribunal de Justiça de São Paulo. A ideia dos autores foi incrementar as análises de jurimetria, habilitando-as para considerar a dinâmica de movimentação das peças jurídicas dentro do fluxo de trabalho do tribunal. Neubauer *et al.* (2022) realizaram uma análise descritiva sobre um log de eventos disponibilizado publicamente²⁷ pela Anvisa (Agência Nacional de Vigilância Sanitária). Nessa análise, alguns algoritmos de descoberta de modelo de processo foram aplicados evidenciando características salientes do processo. Ainda, os autores realizaram uma checagem de conformidade baseada em um modelo de processo normativo derivado da cadeia de valor do processo sob análise, descrita também pela própria agência. Na análise de conformidade, foram reveladas as principais violações a essa cadeia de valor estendida, a origem das violações e como elas impactam na eficiência do trabalho da agência.

Considerações finais

Conforme exposto neste texto, a mineração de processos, por princípio e objetivos, possui grande potencial para ser usada como ferramenta importante na descoberta, entendimento e reflexão sobre o funcionamento dos processos públicos, tanto de apoio, como de entrega de serviços e formulação de políticas públicas. A mineração de processos, sobretudo se empregada com responsabilidade, abre um leque de opções para melhoria de desempenho, auditoria e *compliance*, transparência e participação cidadã, e mudança de visões sobre gestão pública, alinhadas com as novas demandas por

²⁷ <https://dados.gov.br>

transformação e digitalização das diferentes esferas de administração pública no Brasil.

Referências bibliográficas

- BASANYA, R.; OJO, A. (2011). **Mining service integration opportunities towards joined-up government**. In: Proceedings of the 5th International Conference on Theory and Practice of Electronic Governance, pp. 269-272.
- COSTA, H.; RODRIGUES, T. S. (2020). **Mineração de processos aplicada à gestão orçamentária da Marinha do Brasil**. In: Revista Brasileira de Planejamento e Orçamento, 10(1).
- FILGUEIRAS, F.; FLAVIO, C.; PALOTTI, P. (2019), **Digital Transformation and Public Service Delivery in Brazil**. Latin American Policy, 10: 195-219.
- GONZÁLEZ, L.; DELGADO, A. (2021). **Towards compliance requirements modeling and evaluation of e-government inter-organizational collaborative business processes**. *Business Process Management*. 1(2): 3.
- GONZÁLEZ, L.; DELGADO, A.; CANAPARO, J; GAMBETTA, F. (2022). **Evaluation of Compliance Requirements for collaborative business process with process mining and a model of generic compliance controls**. CLEI Electronic Journal, 25(2), 7-1.
- JANOWSKI, T. (2015). **Digital Government Evolution: From Transformation to Contextualization**. Government Information Quarterly, 32 (3), 221-236.
- KALENKOVA, A. A.; AGEEV, A. A.; LOMAZOVA, I. A.; VAN DER AALST, W. M. (2018). **E-government services: Comparing real and expected user behavior**. In: Business Process Management Workshops. Barcelona: BPM 2017 International Workshops, pp. 484-496.
- KUBRAK, K.; MILANI, F.; NOLTE, A.; DUMAS, M (2022). Prescriptive process monitoring: Quo vadis? *PeerJ Computer Science* 8: e1097.
- LARSSON, K.K. (2021). Digitization or equality: When government automation covers some, but not all citizens. *Government Information Quarterly* 38(1): 101547.
- Lei n. 12.527, de 18 de novembro de 2011. (2011). **Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal**. http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm
- Lei n. 14.129, de 29 de março de 2021. (2021). **Dispõe sobre princípios, regras e instrumentos para o Governo Digital e para o aumento da eficiência pública**. http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2021/lei/l14129.htm

-
- MENDES DE ARAUJO, R.; MARQUES PERES, S.; FANTINATO, M.; JACOTO UNGER, A.; RODRIGUES NEUBAUER, T. (2023) **Mineração de Processos como Ferramenta para Promoção da Transparência: Oportunidades e Desafios**. In: *Revista da CGU* 15 (27).
- NEUBAUER, T.R.; ARAUJO, R.M.; FANTINATO, M.; PERES, S.M. (2022) **Transparency promoted by process mining: an exploratory study in a public health product management process**. In: *Anais do X Workshop de Computação Aplicada em Governo Eletrônico (WCGE 2022)*, pp. 37-49.
- PETRUȘEL, R. (2013). A framework for researching public administration decision making processes. *Transylvanian Review of Administrative Sciences*, 9(39): 128-146.
- STEFANINI, A.; ALOINI, D.; DULMIN, R.; MININNO, V. (2016). Linking Diagnostic-Related Groups (DRGs) to their Processes by Process Mining. *HEALTHINF* 5: 438-443.
- REPTA, D.; MOISESCU, M.A.; NAE, A.C.; SACALA, I.S.; DUMITRACHE, I. (2018). Towards document flow discovery in e-government systems. In: *2018 International Symposium on Electronics and Telecommunications (ISETC)*, pp. 1-4.
- SANGIL, M. J. (2020). Heuristics-based process mining on extracted Philippine public procurement event logs. In: *Int. Conf. on Behav. and Social Comput.* (pp. 1–4).
- UNGER, A.J.; SANTOS, J.F. (Neto); TRECENTI, J.; HIROTA, R.; FANTINATO, M.; PERES, S.M. (2021) Process Mining-Enabled Jurimetrics: Analysis of a Brazilian Court's Judicial Performance in the Business Law Processing. In: *Proceedings of 18th International Conference on Artificial Intelligence and Law*, São Paulo (Online).
- VAN DER AALST, W. (2016). **Process Mining: Data science in action**. 2.ed. Berlin: Springer.
- VAN DER AALST, W.; BICHLER, M.; HEINZL, A. (2017). **Responsible data science**. *Business & Information Systems Engineering* 59: 311-313.
- VAN DER AALST, W. (2022). **Process mining**: a 360 degree overview. In: *Process Mining Handbook*. Cham: Springer, pp. 3-34.
- WESKE, M. (2012). **Business Process Management**: Concepts, Languages, Architectures, Verlag. 2.ed. Berlin: Springer.
- WESTPHAL, M.; YOM-TOV, G.B.; PARUSH, A.; CARMELI, N.; ZAMAN, R.; HASSANI, M. (2020). On enabling GDPR compliance in business processes through data-driven solutions. *SN Computer Science*, 1(4).
-

Autores

Sarajane Marques Peres

Professora Associada II na Universidade de São Paulo, na Escola de Artes, Ciências e Humanidades. Trabalha com pesquisa em mineração de processos desde 2014. Foi pesquisadora convidada na Vrije Universiteit Amsterdam e na Utrecht University, Países Baixos, onde aprimorou seu conhecimento na área de mineração de processos. Lattes: <http://lattes.cnpq.br/6265936760089757>

Marcelo Fantinato

Professor Associado III na Universidade de São Paulo, na Escola de Artes, Ciências e Humanidades. Trabalha com pesquisa em mineração de processos desde 2014. Foi pesquisador convidado na Vrije Universiteit Amsterdam e na Utrecht University, Países Baixos, onde aprimorou seu conhecimento na área de mineração de processos. Bolsista CNPq (#312630/2021-2). Lattes: <http://lattes.cnpq.br/8207954538307988>.

Renata Mendes de Araújo

Professora na Faculdade de Computação e Informática e no PPG em Computação Aplicada da Universidade Presbiteriana Mackenzie, do PPG em Sistemas de Informação da EACH-USP e do Doutorado Profissional em Administração Pública da ENAP. Bolsista CNPq (#313210/2019-5). Lattes: <http://lattes.cnpq.br/3589012014320121>.

Fernando Henrique da Silva Costa

Graduado em Sistemas de Informação pelo Centro Universitário de Votuporanga e mestre em Sistemas de Informação pela Universidade de São Paulo, com desenvolvimento de tese na área de inteligência artificial e sistemas de recomendação. Atualmente é doutorando na Universidade de São Paulo na área de mineração de processos. Lattes: <http://lattes.cnpq.br/0523604376837831>.

PROTEÇÃO DOS SEGREDOS INDUSTRIAL E COMERCIAL DA INTELIGÊNCIA ARTIFICIAL E A TRANSPARÊNCIA DA ADMINISTRAÇÃO PÚBLICA

Daniela Copetti Cravo

Introdução

Na área da tecnologia da informação, os segredos industrial e comercial são, frequentemente, utilizados para fins de preservar a competitividade e a vantagem concorrencial das empresas. Ocorre, todavia, que uma utilização sem critérios dos segredos industrial e comercial pode prejudicar a transparência e a divulgação de informações necessárias para o interesse público.

No contexto da justiça criminal, os obstáculos do segredo comercial à transparência algorítmica são especialmente problemáticos, uma vez que a liberdade individual está em jogo (BRAUNEIS e GOODMAN 2018). Nessa temática, algumas experiências negativas deixaram lições e reflexões importantes.

A respeito, cita-se o caso, nos Estados Unidos, de réu em processo penal que foi condenado a seis anos de prisão, em decisão que utilizou o sistema COMPAS. Ao solicitar informações relacionadas a como o COMPAS pesou determinadas variáveis de entrada e como essas informações foram calculadas para sua pontuação de risco final a respeito da probabilidade de sua reincidência, o réu teve seu pedido negado, sob alegação de que tais informações estariam protegidas pelo segredo comercial (MOORE 2017).

Tal problemática é vivenciada em outros campos da esfera pública quando do uso de sistemas de inteligência artificial (IA) por

órgãos públicos. Pesquisas revelaram que nem sempre é possível identificar o uso desses sistemas, já que muitos órgãos públicos alegam não poder compartilhar tais informações por causa do segredo comercial (WRAY 2022).

Em um levantamento feito por Brauneis e Goodman (2018), após o envio de 42 pedidos de informações sobre programas de algoritmos preditivos no poder público, verificou-se que em quase todos os casos não foram fornecidas informações. Um problema apontado pelos autores nessa experiência foi a recorrência de afirmações excessivamente amplas de sigilo comercial.

Desta forma, muito mais que uma preocupação meramente teórica, a criação de uma opacidade ainda maior aos sistemas de inteligência artificial por meio dos segredos industrial e comercial é algo que já tem ocorrido na prática. Por tal razão, é essencial adotar uma abordagem equilibrada na interpretação das normas relacionadas aos segredos industriais e comerciais quanto à implementação da proteção e aos seus limites.

Com efeito, o objetivo do presente artigo é, justamente, apresentar algumas propostas interpretativas de equilíbrio entre a transparência pública no uso de sistemas de inteligência artificial e a observância dos segredos industrial e comercial.

Possível Tensão entre Transparência Pública e Proteção dos Segredos Industrial e Comercial Relacionados aos Detalhes Técnicos da Inteligência Artificial

A administração pública precisa ser transparente, no sentido de que o processo de tomada de decisão deve ser justificável e, em certa medida, compreensível pelo público em geral. Além disso, os cidadãos geralmente têm direito a explicações sobre decisões que os impactam (SUPREME AUDIT INSTITUTIONS OF FINLAND, GERMANY, THE NETHERLANDS, NORWAY AND THE UK 2023). Por tal razão, a opacidade e a complexidade dos sistemas de Inteligência Artificial não devem ser usadas para tornar as demandas por transparência insignificantes (MYLLY 2023).

Lembre-se que a transparência na administração pública é um pilar fundamental da democracia e do Estado de Direito. Essa garante

que as decisões governamentais sejam baseadas em motivos claros e justificáveis, promove a responsabilidade dos gestores públicos e capacita os cidadãos a participarem ativamente na construção de um governo mais eficaz e responsivo às necessidades da sociedade. Portanto, é essencial que o poder público seja exercido “em público”²⁸, de acordo com os princípios de transparência e prestação de contas.

A essencialidade da transparência administrativa também está ligada ao reconhecimento do direito fundamental a uma boa administração²⁹. Os cidadãos têm o direito de receber um serviço público de qualidade, e a transparência é uma ferramenta vital para garantir que a administração pública atenda a essas expectativas.

Além disso, a transparência na administração pública promove a participação popular e a liberdade de expressão. Quando as informações são acessíveis e compreensíveis, os cidadãos podem se envolver mais ativamente na tomada de decisões e no debate público. Isso fortalece a democracia, permitindo que as vozes da sociedade sejam ouvidas e levadas em consideração no processo político.

Pois bem, considerando a importância da transparência pública, por um lado, e a crescente tendência de uso de sistemas de inteligência artificial pelo poder público³⁰, por outro, como solucionar a possível tensão existente entre as regras relacionadas à transparência pública com aquelas relacionadas aos segredos industrial e comercial dos detalhes técnicos da Inteligência Artificial? Para bem endereçar esta questão, passa-se a analisar a temática dos segredos industrial e comercial, o que esses abrangem e como devem ser tutelados.

²⁸ Segundo BOBBIO (1986), a democracia é o governo do “poder público em público”.

²⁹ O direito fundamental à boa administração pode ser entendido como o atendimento dos “objetivos fundamentais da Constituição por meio de uma administração que concilie os custos orçamentários com os interesses e necessidade do administrado, ou seja, a pessoa humana, razão de ser do Estado” (ISMAIL FILHO, 2018).

³⁰ Sobre iniciativas de transparência, citam-se as iniciativas de Amsterdã e Helsinque, que laçaram registros de Inteligência Artificial para detalhar como cada governo municipal usa algoritmos para fornecer serviços públicos. Nesse registro, são listados os algoritmos utilizados, bem como os conjuntos de dados usados para treinar um modelo, a descrição de uso do algoritmo, além das possíveis tendências ou riscos. Ver: <https://www.itu.int/en/myitu/News/2020/09/30/07/41/Helsinki-Amsterdam-AI-registers-city-systems-Cities-Today>.

Observância dos Segredos Industrial e Comercial

No direito brasileiro não há uma definição precisa do que seja segredo de empresa (SILVEIRA 2001), o qual abarca as espécies segredo comercial e segredo industrial³¹. Na verdade, consoante expõe Fekete (2003), o sistema brasileiro de proteção do segredo de empresa é baseado nas normas de repressão à concorrência desleal.

Pela (2018) salienta que, apesar da inegável relevância do tema para a prática empresária e para a estruturação do sistema de propriedade intelectual, o campo ainda é pouco explorado no Brasil. A violação do segredo de empresa está, no Brasil, relacionada à prática de concorrência desleal (art. 195, XI e XII da Lei 9.279/1996), a qual é qualificada como um crime na legislação brasileira. Apenas de forma secundária são reconhecidos os efeitos civis de práticas dessa natureza (PELA 2018).

Sobre a definição de concorrência desleal, Parente (2005) destaca que, muito embora não se saiba exatamente como defini-la, é senso comum caracterizá-la quando a competição entre concorrentes é exercida de maneira contrária às práticas usuais dos empresários e de forma imoral, abusiva e desonesta³².

Complementa, ainda, a autora, que os doutrinadores costumam trazer exemplos do que seria segredo de empresa. Podem ser apontados os seguintes: os dados sobre clientes, meios e processos mediante os quais os empresários atraem a clientela, aumentando-a e que, devido à importância estratégica, devem ser mantidos em sigilo (PARENT, 2005). Para Silveira (1981), o segredo industrial seria um bem imaterial de exclusividade imperfeita, sendo elemento do estabelecimento empresarial.

Nos Estados Unidos, a título de comparação, o segredo de empresa é mais consolidado, tendo um crescimento significativo na sua importância, inclusive com um *boom* de litígios nessa temática

³¹ Tendo em vista que a expressão “segredo de empresa” é um gênero que acaba abrangendo as demais espécies, como os segredos comercial e industrial (BARBOSA 2013), optou-se por utilizar o gênero “segredo de empresa”, no presente artigo.

³² Ver: Ascarelli (1957).

(ALMELING, 2012). Já na União Europeia, o tema foi objeto da Diretiva 94/3/1996, que reconhece que as informações são a moeda da economia do conhecimento e, como tal, uma vez qualificadas para proteção de segredos comerciais, merecem proteção em termos legais em toda a União Europeia. Tal harmonização serve, ademais, para promover o *Single Market* europeu (FALCE 2015).

Para Labrunie (2000), o segredo de negócio é, sem dúvida, um fato e não um direito em si. Com efeito, não se vislumbria, no direito pátrio, nem a possibilidade de direito de propriedade, nem proteção da posse, uma vez que o cerne recai sobre o segredo e o conhecimento. Assim, na esteira do mencionado autor, o legislador pátrio teria entendido que o segredo é um *plus*, uma vantagem mercadológica, de natureza concorrencial e, por isso, inserida no contexto da repressão à concorrência desleal.

Sobre o que deve ser entendido como segredo, Fekete (2003) pontua que qualquer informação da empresa, decorrente de investimento de tempo, dinheiro e trabalho, pode ser protegida pela legislação do segredo de negócio, desde que observadas as condições legais. O primeiro requisito para o empresário fazer jus à tutela legal diz respeito a um conhecimento que deve ser mantido sob reserva ou sigilo, o que não significa o total “não acesso”, já que esse pode ocorrer para fins de desempenho das atividades rotineiras à empresa.

Ao contrário, ficam excluídas da proteção as informações de conhecimento público, fulcro no art. 195, XI da LPI. Situações comportamentais do detentor do segredo que demonstram a sua intenção de não tornar públicas tais informações manifestam-se pela celebração de termos de confidencialidade com os funcionários contratados, sendo possível que tais pactos sejam firmados, após a ruptura do contrato de trabalho³³, bem como o não pedido de

³³ CONTRATO DE TRABALHO (EM GERAL) – CLÁUSULA ILEGAL – DANOS MATERIAIS – INDENIZAÇÃO – INEXISTÊNCIA – Não afronta o art. 5º, inc. XIII da Constituição Federal de 1988 cláusula contratual firmada por empregado, após ruptura do contrato de trabalho, comprometendo-se a não prestar serviços à empresa concorrente, quer como empregado, quer como autônomo. Inexistência de erro ou coação a anular o pactuado. Não há ilegalidade a macular ao pactuado e nem danos materiais decorrentes da limitação expressamente aceita. Impende aqui, invocar-se também o princípio da boa-fé, presente em todos os atos da vida civil e pressuposto deles, mormente quando em ajuste, estão pessoas capazes, de mediano conhecimento

patente, uma vez que esses acabam sendo publicados na Revista da Propriedade Industrial (RPI) do Instituto Nacional da Propriedade Industrial (INPI). Ou seja, para fins de tutela, é indispensável que as informações tenham sido objeto de precauções razoáveis, as quais demonstram a intenção de manter a informação em segredo (FEKETE, 2003).

Para tanto, sempre será necessária uma análise feita caso a caso. Ademais, o bem jurídico tutelado pelo segredo de empresa deve possuir valor econômico, não podendo o seu conteúdo ser irrelevante no mercado. Trata-se de algo que configure uma vantagem sobre os concorrentes e que tenha aplicabilidade ao negócio da empresa (FEKETE, 2003).

Nesse contexto, indaga-se: como evitar que a proteção dos segredos industrial e comercial relacionada aos sistemas de inteligência artificial crie uma opacidade indevida à transparência pública?

Possíveis Soluções

Conforme destacam Noto e Sappa (2020), os segredos industrial e comercial são muitas vezes considerados como uma opção mais viável do que as patentes, em particular na área da tecnologia da informação. Por causa disso, é possível que a divulgação de informações e a transparência voltada para o interesse público sejam mais marginais do que o adequado. Com efeito, segundo os mencionados autores, a única maneira de se recuperar dessa tendência é por meio de uma leitura equilibrada das normas dos segredos industrial e comercial quanto à implementação da proteção e aos seus limites.

No âmbito europeu, e tendo como base a proposta de Regulamento da Inteligência Artificial, Mylly (2023) aponta que os segredos comerciais não deveriam criar tantos obstáculos à transparência algorítmica como poderia se supor inicialmente. O

jurídico e alto nível profissional, como é o caso das partes envolvidas no Termo de Confidencialidade e Compromissos Recíprocos.’ (BRASIL.TRT 2ª Reg. RO 20020208981 _ (20020534563) 5ª T. Rel. Juíza Rita Maria Silvestre _ DOESP 30.08.2002).

primeiro fundamento para essa conclusão é que, mesmo que uma explicação não seja completa, ainda assim essa pode ser suficiente para o propósito de alcançar um controle significativo.

Ou seja, é possível que sejam dadas informações que sejam significativas à transparência pública, sem que tais informações revelem ou coloquem em vulnerabilidade o segredo comercial ou industrial. Transparência não significa necessariamente “transparência total”, de forma que o necessário é que a explicação seja adequada aos fins visados, ainda que essa não seja necessariamente completa (MYLLY 2023).

Impende apontar também que o desenvolvimento de sistemas de inteligência artificial pode depender de uma grande quantidade de componentes de ciência básica e módulos de *software* pré-existentes. Também há o desenvolvimento de sistemas pautados em *standards*, os quais, para cumprir o seu papel, devem ser abertos. Para mais, os sistemas e inteligência artificial contêm muitos componentes que utilizam dados coletados de fontes disponíveis publicamente.

Nesses casos, parece razoável concluir que situações tais como as narradas acima talvez não cumpram os requisitos para fins de proteção dos segredos industrial e comercial. Tais características do desenvolvimento de sistemas de inteligência artificial aparentemente reduzem a extensão da proteção dos segredos industrial e comercial. Todavia, a maneira específica como esses componentes geralmente conhecidos são reunidos pode vir a ser considerada abarcada no segredo (MYLLY 2023).

Para mais, entende-se que as negativas de acesso a documentos ou informações devem explicar claramente de que forma, específica e objetivamente, o acesso aos documentos poderia prejudicar o interesse protegido pelos segredos industrial e comercial. A respeito, lembre-se que a negativa de acesso às informações deve ser sempre fundamentada, sob pena de responsabilização nos termos do artigo 7º, §4º, da LAI (Lei de Acesso à Informação).

Pontua-se que o detentor dos segredos industrial e comercial pode ter apenas uma alegação ou uma crença quanto a uma possível

proteção. No entanto, tal suposição não gera um direito absoluto quanto ao sigilo³⁴.

Nesse sentido, pondera-se que mesmo que a LAI preveja no seu artigo 22 o dever de guarda do segredo industrial ou comercial, ainda assim, tal dever só tem potencial incidência quando a divulgação colocar em risco a atividade empresarial³⁵.

É preciso, pois, que haja um equilíbrio de interesses, mesmo quando aplicável a proteção à informação da empresa. Insta pontuar que nem todos os detalhes técnicos da Inteligência Artificial irão necessariamente se qualificar como segredo, de forma que uma negativa geral a respeito pode configurar um exercício inadmissível de posições jurídicas (ou abuso do direito)³⁶.

Ademais, é imperioso registrar que o acesso ao código-fonte geralmente não será necessário para se alcançar uma transparência significativa. Desta forma, no que as entidades públicas deveriam estar focadas é em realizar a concepção, aquisição e implementação de processos algorítmicos de forma ponderada e transparente (BRAUNEIS e GOODMAN 2018).

Outrossim, nas licitações e contratações públicas deve ser exigido dos fornecedores que esses criem e entreguem registros que expliquem as principais decisões e os esforços de validação, sem necessariamente divulgar fórmulas ou algoritmos precisos (BRAUNEIS e GOODMAN 2018).

Tal obrigação deve ser pensada desde a fase preparatória da contratação de sistemas de inteligência artificial e prevista na minuta do contrato. Cabe dizer que, desde o estudo técnico preliminar³⁷

³⁴ Nesse ponto, destaca-se a importância de os órgãos públicos compreenderem os conceitos relacionados aos segredos industrial e comercial para fins de poderem tomar decisões corretas a respeito (MYLLY 2023).

³⁵ Critério utilizado pela CGU em casos de restrição do acesso à informação conforme o seguinte precedente: 99926000080201785. Disponível em: <http://buscaprecedentes.cgu.gov.br/busca/dados/Precedente/99926000080201785_CGU.pdf>.

³⁶ Ver: MENEZES CORDEIRO (1997).

³⁷ Por analogia, cita-se conclusão semelhante no que toca aos dados pessoais (FORTINI e SOARES 2023): “O ETP, ao identificar a solução, também deve atentar para a questão do tratamento dos dados pessoais inclusive avaliando alternativas que possam minimizar riscos a isso relacionados.”

(artigo 6º, inciso XX, da Lei 14.133/2021), já devem ser analisadas as alternativas e soluções para se promover a transparência desde a concepção.

Considerações Finais

A administração pública deve operar de maneira transparente, garantindo que o processo de tomada de decisão seja justificável e compreensível para o público em geral. Dessa forma, a proteção aos segredos industrial e comercial dos aspectos técnicos da inteligência artificial não pode ser utilizada para criar, de forma indevida, barreiras à transparência desejada.

Conforme exposto ao longo do artigo, a transparência não exige necessariamente uma divulgação completa, mas sim explicações adequadas para atender aos objetivos desejados. Também deve ser levado em consideração o fato de que nem todos os detalhes técnicos da inteligência artificial qualificam-se como segredos.

Ademais, as negativas de acesso a informações devem ser devidamente fundamentadas, explicando de maneira clara como o acesso poderia prejudicar o interesse protegido pelos segredos industriais e comerciais. Além disso, é importante que as entidades públicas compreendam os conceitos relacionados a esses segredos para tomar decisões informadas.

Por fim, nas licitações e contratações públicas, os fornecedores devem ter a obrigação de fornecer registros explicativos sobre o sistema de inteligência artificial, sem necessariamente divulgar fórmulas ou algoritmos precisos. Da mesma forma, as entidades públicas devem se concentrar em implementar processos algorítmicos da maneira mais ponderada e transparente possível, promovendo assim uma maior compreensão das principais decisões tomadas.

Referências Bibliográficas

ALMELING, D.S. (2012) Seven Reasons why Trade Secrets are Increasingly Important, 27 *Berkeley Technology Law Journal*.
ASCARELLI, T. (1957) **Teoria dela concorrenza e dei beni immateriali. Lezioni di dirritto industriale**. 2.ed. Giuffrè: Milano, p. 179-213.

-
- BARBOSA, D.B. (2013) **Tratado de Propriedade Intelectual**. Tomo I. Rio de Janeiro: Lumen Juris.
- BOBBIO, N. (1986). **O Futuro da Democracia**. Rio de Janeiro: Paz e Terra.
- BRASIL.TRT 2ª Reg. RO 20020208981 – (20020534563) 5ª T. Rel. Juíza Rita Maria Silvestre – DOESP 30.08.2002.
- BRAUNEIS, R.; GOODMAN, E. (2018) Algorithmic Transparency for the Smart City. 20 *Yale J. of Law & Tech*. 103.
- FALCE, V. (2015) Trade Secrets – Looking for (Full) Harmonization in the Innovation Union. IIC – *International Review of Intellectual Property and Competition Law*, v. 46, p. 964.
- FEKETE, E.K. (2003) **O regime jurídico do segredo de indústria e comércio no Direito Brasileiro**. Forense: Rio de Janeiro.
- FORTINI, C.; SOARES, A. (2023) Dados pessoais nas contratações públicas: cuidados na fase preparatória. Disponível em : < <https://www.conjur.com.br/2023-jun-01/interesse-publico-dados-pessoais-contratacoes-publicas-fase-preparatoria> >. Acesso em 8 de set. de 2023.
- ISMAIL, S. (Fº) (2018). Boa administração: um direito fundamental a ser efetivado em prol de uma gestão pública eficiente. *Revista de Direito Administrativo* 277(3): 105–137.
- LABRUNIE, J. (2000) **Proteção ao Segredo de Negócio. Direito Empresarial Contemporâneo**. São Paulo: Juarez de Oliveira, p. 93-94.
- MENEZES CORDEIRO (1997). **Da Boa Fé do Direito Civil**. Coimbra: Almedina.
- MOORE, T. (2017) *Trade Secrets and Algorithms as Barriers to Social Justice*. CDT Free Expression Fellow.
- MYLLY, U.M. (2023) *Transparent AI? Navigating Between Rules on Trade Secrets and Access to Information*. IIC 54, 1013–1043.
- NOTO, L.D.G; SAPPA, C. (2020) The internet of things (IoT) at the intersection of data protection and trade secrets. non-conventional paths to counter data appropriation and empower consumers. *Eur J Consum Law/Revue européenne de droit de la consommation* 3:419–458.
- PARENTE, F. (2005) Concorrência Desleal e Segredos de Fábrica e de Negócios – Análise do art. 195 XI da Lei de Propriedade Industrial.

Revista de Direito Mercantil Industrial, Econômico e Financeiro. 139, p. 182.

PELA, J.K. (2018) The Brazilian Regulation of Trade Secrets: a proposal for its review. *Gewerblicher Rechtsschutz und Urheberrecht – Internationaler Teil*, v. 6.

SILVEIRA, J.M. (2001) A Proteção Jurídica dos Segredos Industriais e de Negócio. *Revista de Direito Mercantil, Industrial, Econômico e Financeiro* 40(121), p. 152.

SILVEIRA, N. (1981) Limites Convencionais à Concorrência. *Revista de Direito Mercantil Industrial Econômico e Financeiro*, Nova Serie, 20, p. 48.

SUPREME AUDIT INSTITUTIONS OF FINLAND, GERMANY, THE NETHERLANDS, NORWAY AND THE UK (2023). Auditing machine learning algorithm. *A white paper for public auditors*.

WRAY, S. (2022) Cities confront the challenges of algorithm transparency. Disponível em: < <https://cities-today.com/cities-confront-the-challenges-of-algorithm-transparency/>>. Acesso em 8 de set. de 2023.

Autora

Daniela Copetti Cravo

Procuradora do Município de Porto Alegre. Doutora e Mestre em Direito pela UFRGS. Pós-Doutorado no Departamento de Direito Público e Filosofia do Direito da Faculdade de Direito da UFRGS (2019 - 2020).

RECONHECIMENTO FACIAL EM ESPAÇOS PÚBLICOS: um risco para os regimes democráticos

**Rodrigo Brandão
Thiago Gomes Marcilio**

Introdução

A aplicação de reconhecimento facial na segurança pública tem sido ampliada. O reconhecimento facial é uma técnica que analisa as características faciais presentes em imagens da face, como os olhos, sobrancelhas, nariz, lábios e formato do queixo.

O encontro de uma amplitude de dados coletados, uma capacidade de processamento computacional, o uso e treinamento de algoritmos, a criação de uma malha de câmeras de vigilância nas grandes metrópoles e um momento político de retrocesso democrático que torna a administração pública mais permeável ao uso destas tecnologias.

Essa combinação cria riscos ao Direito de Protesto, direito exercido de forma excepcional, mas cujo potencial comprometimento pode consolidar uma trilha de restrições e ameaças a outros direitos e à própria capacidade do Estado de garantir o cumprimento de suas estratégias de segurança.

Ondas Democráticas

A história das democracias no século XX pode ser organizada em três “ondas democráticas”, uma abordagem estruturada apresentada por Samuel P. Huntington. A primeira onda foi longa,

começando na década de 1820 e durando até 1922, quando o Fascismo na Itália e o Nazismo na Alemanha a interromperam. A segunda veio logo após a Segunda Guerra Mundial, quando 36 governos aderiram a esse regime.

No entanto, quase como um pêndulo, uma segunda onda reversa pôde ser observada, e mais uma vez as democracias foram restringidas por ditaduras, muitas das quais foram apoiadas financeira ou militarmente pelos Estados Unidos ou pela União Soviética. A terceira onda democrática ocorreu entre 1974 e 1990. As democracias se espalharam pelo globo, alcançando 30 países em 15 anos, e na época a questão era se esse regime se expandiria indeterminadamente (HUNTINGTON 1991, p. 21).

Huntington (1991, p.14-15) indicou cinco principais fatores que contribuíram para essa terceira onda e permitiriam que ela se mantivesse forte, que foram: (i) os problemas de legitimidade enfrentados pelos governos autoritários devido a falhas militares ou econômicas e uma falta geral de sucessos; (ii) o crescimento econômico, que expandiu a classe média e melhorou a qualidade de vida; (iii) a mudança da Igreja Católica para defender uma agenda mais progressista, incluindo a luta contra o autoritarismo; (iv) Mudanças de atores externos; (v) e acumulação de elementos que tenderam a uma abordagem democrática.

O autor também indica que (vi) a maior parte dos governos protestantes havia se convertido em regimes democráticos e, até 1990, o mesmo ocorreu com as nações católicas. (vii) Durante essa onda, a Comunidade Europeia nasceu e os benefícios econômicos se tornaram atrativos para o Continente, o que transformou o estabelecimento de uma democracia em um passo importante para a estabilidade econômica. Grécia, Espanha, Portugal e Turquia foram exemplos dessa transição de governos autoritários para democráticos.

O fim da União Soviética levou a democracia a seus herdeiros. Se a Rússia se tornasse uma democracia, isso seria um passo importante para um regime democrático. Em 1991, isso era uma esperança, mas em 2022, após 22 anos de governo de Putin, a Rússia se vê novamente em um sistema autocrático, rimando com a Rússia czarista e os presidentes da União Soviética.

A doutrina dos Estados Unidos de promoção da democracia também foi um fator importante, na visão dele, uma vez que o país oferecia benefícios comerciais para novas democracias. No entanto, o deterioramento dessa própria doutrina poderia oferecer um alerta.

Se as pessoas ao redor do mundo passarem a ver os Estados Unidos como um poder enfraquecido, assolado pela estagnação política, ineficiência econômica e caos social, seus fracassos percebidos inevitavelmente serão vistos como os fracassos da democracia, e o apelo mundial da democracia diminuirá. (tradução livre) (HUNTINGTON, 1991, p. 16)

Um elemento imprevisível deve ser adicionado, a disseminação da Internet e o acesso individualizado em massa às informações por meio das redes sociais.³⁸ Como resultado, provavelmente estamos enfrentando um encolhimento do espírito democrático, uma Terceira Onda Reversa (HUNTINGTON 1991, p. 7).

O ponto é que os elementos trazidos à luz por Huntington estão desvanecendo. O mundo enfrentou duas crises econômicas, uma em 2009, causada pelas hipotecas *subprime* americanas, e outra em 2020, devido à pandemia. (HUNTINGTON 1991, p. 7)

Um sentimento de insatisfação com a eficiência institucional também gerou desconfiança nos sistemas democráticos, o que culminou na "fragilidade dos valores democráticos entre grupos de elite chave e o público em geral" (HUNTINGTON, 1991, p. 7). Um exemplo disso é a luta anticorrupção no Brasil, que foi iniciada devido a escândalos de suborno no Executivo e resultou no *impeachment* da ex-presidente Dilma Rousseff.

No contexto da invasão ucraniana, a mídia concluiu que os EUA encolheram e não podem mais garantir a segurança da Europa

³⁸ No passado, notícias falsas foram usadas para propaganda política e militar e até mesmo para criar narrativas no setor privado, a fim de tornar um produto viável aos olhos dos consumidores. No entanto, no passado, essas informações falsas precisavam ter a mesma escala dos meios de comunicação tradicionais. A sobreposição da internet, telefones celulares e doses diárias de informações falsas têm representado uma ameaça aos regimes democráticos com gastos menores e resultados mais significativos.

(BBC 2022). A aparência de ineficiência democrática pode levar à uma legislação mais propensa a aceitar a aplicação de ferramentas que podem ter usos autoritários e que relativizam direitos fundamentais, especialmente tecnologias que permitem uma identificação mais fácil de dissidentes, opositores e daqueles que denunciem violações dos direitos humanos.

Um dos efeitos desse retrocesso democrático seria uma abordagem tecnocrática, por meio da qual o uso da tecnologia tornaria possível manipular comunicações em massa (HUNTINGTON, 1991, p. 10) e, noutra ponta, diminuir a capacidade de articulação da população para protestar publicamente, se reunir e expressar seus desejos.

Direito de Protestar

Em diversos momentos políticos, os protestos foram utilizados como ferramenta de expressão da insatisfação de uma população com relação a um determinado evento, decisão ou abordagem política. Exemplos são a Marcha sobre Washington (1963); os protestos de 68 na França (1968); o movimento "Diretas Já" (1984) no Brasil; a eleição presidencial no Irã (2009); a Primavera Árabe (2010); o *Occupy Wall Street* (2010); os protestos contra a corrupção no Brasil (2013); os protestos em Hong Kong (2019-2020) e o movimento *Black Lives Matter*, que ocorreu globalmente (2020).

Protestar é uma das maneiras de participar de um regime democrático e criar um sistema ativo que garanta que o povo legitime seus líderes e exerça a autodeterminação política na “formulação de normas jurídicas por aqueles mesmos que devem obedecê-las, com exclusão de toda influência estranha” (GARCIA 2003, p. 16), que é a essência da democracia (p. 21). As manifestações encontrarão apoio legal no direito de se associar livremente e se reunir em espaços públicos e no exercício da liberdade de expressão. O Artigo 19 e 20 da Declaração Universal dos Direitos Humanos também estabelecem isso (ONU 1948).

Artigo 19 Todos têm o direito à liberdade de opinião e expressão; esse direito inclui a liberdade de ter opiniões sem interferência e de buscar, receber e divulgar informações e ideias por qualquer meio e independentemente de fronteiras.

Artigo 20 Todos têm o direito à liberdade de reunião pacífica e associação. Ninguém pode ser obrigado a pertencer a uma associação.

A própria natureza dessa ferramenta reside no poder que está nas multidões, nas quais cada indivíduo se torna parte do discurso e decide deixar de lado sua individualidade para promover um interesse compartilhado ou uma lista completa de valores, crenças e demandas. Nosso regime aceita essa forma de expressão extra-institucional. Os protestos são ações excepcionais, mas a identificação de uma ameaça a esta prática pode criar riscos para outros direitos.

Reconhecimento facial amplo

Uma lista de casos e cenários arriscados pode ser apontada. Em 2020, o Departamento de Polícia de Nova York utilizou o reconhecimento facial para localizar e prender o ativista do *Black Lives Matter*, Derrick "Dwreck" Ingram. Por cinco horas, seu apartamento foi sitiado e os policiais tentaram interrogá-lo, mesmo sem a presença de seu advogado.

A acusação referia-se ao uso de um megafone e gritos contra um policial em 14 de junho, um mês e meio antes da operação policial. Um relatório de reconhecimento facial foi visto na mão de um policial e "posteriormente, foram criados cartazes de 'Procurado' com fotos privadas de Ingram do Instagram e esses cartazes foram colados no bairro de Dwreck e nas redes sociais do NYPD." (ANISTIA INTERNACIONAL 2022a).

Em 2018 e 2019, a Nelson Management Group, proprietária do prédio residencial Atlantic Plaza Towers, em Brownsville, utilizou sistemas de reconhecimento facial em suas câmeras de segurança principalmente para multar os moradores que cometeram infrações menores, como "não separar recicláveis" (ANISTIA INTERNACIONAL

2022a) e para punir aqueles que distribuíam panfletos com alertas sobre a implementação do reconhecimento facial.

Esses "inquilinos receberam cópias coloridas das imagens das câmeras de vigilância, onde estavam na portaria, com seus números de apartamento e data e hora escritos nas fotos" (ANISTIA INTERNACIONAL 2022a). Após uma longa batalha com os proprietários, o uso do sistema foi interrompido.

Em 2019, seguindo os usos anteriores do reconhecimento facial, o Ministério do Interior da Índia propôs o Sistema de Reconhecimento Facial da Índia, um banco de dados centralizado para todo o país. Um Centro de Comando e Controle (CCC) foi construído em Hyderabad para processar 600.000 câmeras. A escolha da cidade não foi por acaso, uma vez que Hitec City ou Cyberabad, um microdistrito, está localizado na região. Em apenas dois bairros, Kala Pathar e Kishan Bagh, a área de vigilância mapeada era de quase 1,06 milhão de metros quadrados. Policiais também tiraram fotos e fizeram solicitações aleatórias para coletar impressões faciais e digitais (ANISTIA INTERNACIONAL 2022b).

Em 2021, a polícia de Moscou deteve centenas de manifestantes que apoiavam o candidato da oposição, Aleksei Navalny. Na época, surgiu um temor generalizado sobre o risco do reconhecimento facial (ANISTIA INTERNACIONAL 2021). Se essa tecnologia tivesse sido usada na época, poderia ter sido usada agora, durante os protestos contra a invasão russo-ucraniana, nos quais mais de 13.000 pessoas foram presas até 6 de março, segundo a Human Rights Watch (2022).

A América Latina também é um local potencial para a coleta massiva de dados. Em 3 de março, um grupo de organizações de direitos humanos entrou com um processo de emergência no Tribunal de Justiça do Estado de São Paulo, solicitando a proibição de qualquer forma de coleta de dados, a suspensão da captura de dados biométricos ou a instalação de novas câmeras. A implementação foi suspensa pela primeira instância do tribunal em 22 de março de 2022, uma vez que a explicação técnica não foi considerada suficiente. (TRIBUNAL DE JUSTIÇA DE SÃO PAULO, processo de 1010667-97.2022.8.26.0053). Foram apontadas também violações à Lei

Brasileira de Proteção de Dados, uma legislação federal, e a falta de precisão do sistema, o que poderia levar a prisões injustas.

O reconhecimento facial nos aeroportos de São Paulo e Rio de Janeiro ainda está em operação.³⁹ Essa forma de coleta massiva de dados não é nova, como o caso da coleta de dados no Aeroporto de Heathrow em 2003, e a violação da Lei de Proteção de Informações Biométricas de Illinois (BIPA) por gigantes da tecnologia, que coletaram e processaram fotografias disponíveis na Internet para criar um banco de dados facial.

Essa tecnologia representa uma ameaça aos direitos dos cidadãos, uma vez que possui um grande poder de causar violações dos direitos humanos, ao coletar, seguir, rastrear e invadir a privacidade das pessoas. Também foi responsável por prisões injustas nos EUA (THE NEW YORK TIMES 2020), Argentina⁴⁰ e Brasil⁴¹, incluindo uma falha em 2023 que levou à detenção de uma mulher grávida de oito meses por engano⁴², e pela vigilância de minorias étnicas na Itália⁴³. Além disso, poderia operacionalizar teorias eugênicas ou facilitar caçadas e genocídios.

³⁹ BRASIL. Ministério da Casa Civil. Aeroportos de Congonhas (SP) e Santos Dumont (RJ) adotam embarque facial biométrico 100% digital. Publicado em 10/08/2022 11h43. Atualizado em 31/10/2022 11h30. Disponível em: <online> Acesso em: 01 set. 2023.

⁴⁰ WIRED. Un escándalo en Buenos Aires revela los peligros del reconocimiento facial. 15 de setembro de 2023. Disponível em: <https://es.wired.com/articulos/escandalo-en-buenos-aires-revela-los-peligros-del-reconocimiento-facial>. Acesso em: 23 de setembro de 2023.

⁴¹ G1 RIO. Sistema de reconhecimento facial da PM do RJ falha, e mulher é detida por engano. 11 de julho de 2019. Disponível em: <https://g1.globo.com/rj/rio-de-janeiro/noticia/2019/07/11/sistema-de-reconhecimento-facial-da-pm-do-rj-falha-e-mulher-e-detida-por-engano.ghml>. Acesso em: 23 set. 2023.

⁴² UOL. Grávida de 8 meses é presa após erro em reconhecimento facial. 22 de setembro de 2023. Disponível em: <https://gizmodo.uol.com.br/gravida-de-8-meses-e-presa-apos-erro-de-software-de-reconhecimento-facial/>. Acesso em: 22 set. 2023

⁴³ WIRED. Perché Como è diventata una delle prime città in Italia a usare il riconoscimento facciale. Disponível em: <https://www.wired.it/internet/regole/2020/06/09/riconoscimento-facciale-come/>. Acesso em: 08 set. 2023.

Considere a tecnologia de impressões digitais, que teve origem na África do Sul para a vigilância, monitoramento e mapeamento da população negra. Essa mesma tecnologia foi posteriormente aprimorada, levando eventualmente à eugenia e à Alemanha Nazista. E agora, todos estamos lidando com a perspectiva do reconhecimento facial sendo usado para nos vigiar a todos. (MILLER 2022)

Os cenários apresentados neste artigo oferecem uma visão geral de como esses sistemas, em pleno funcionamento e sem vieses, podem ser perigosos para os direitos fundamentais, especialmente aqueles que visam garantir a sustentabilidade interna dos regimes democráticos. No entanto, há outros direitos e grupos vulneráveis que também podem ser prejudicados.

Aplicações e falhas contra grupos específicos

Na cidade de São Paulo, a Prefeitura chamou empresas privadas para doar ou emprestar tecnologias para a Segurança Pública, com o objetivo de verificar a viabilidade dessas técnicas, melhorar a eficiência e reduzir gastos. "*Tasers*", "*Body Worn*" e *drones* com reconhecimento facial foram solicitados no chamamento público. O uso de *drones* criaria um banco de dados da população viciada em drogas, analisaria a migração dessas pessoas na cidade e evitaria sua concentração em espaços públicos (PREFEITURA DE SÃO PAULO 2020, p. 32) Isso pode ser considerado um "tecno-higienismo". Em um exercício especulativo, armas não letais também poderiam ser usadas em tais *drones* não tripulados.

Esse fenômeno existe em todas as áreas urbanas, mas essa não é a abordagem recomendada para lidar com aqueles que são viciados em drogas. Rastreá-los é o primeiro passo para entender quais serviços sociais e públicos eles buscam e, eventualmente, tornar esses serviços mais difíceis de acessar, o que tornaria sua sobrevivência extremamente difícil. Mapear seus movimentos também pode ser uma maneira de identificar quais serviços de apoio devem ser fechados. Também é importante lembrar que a população em situação de rua enfrenta grandes dificuldades para acessar

instituições, incluindo tribunais e defensores públicos, o que torna ainda mais difícil sua autoproteção.

Outro tópico importante em São Paulo é o direito à moradia. Devido à concentração econômica na região Sudeste, principalmente em São Paulo e Rio de Janeiro, a reserva de mão de obra se tornou a maior do país, o que gera uma rotação em termos de emprego. Aqueles que estão trabalhando podem se dar ao luxo de alugar uma casa, mas eventualmente podem ficar desempregados e viver em condições precárias de habitação.

Existem também movimentos sociais cujos membros não têm capacidade financeira para acessar moradias, devido à falta de oferta de habitação pública. Relatórios mostram que há pelo menos 20 mil famílias aguardando por uma casa ou apartamento público ou vivendo em situações precárias na cidade de São Paulo (G1, 2021, s/n) Na mesma cidade, existem pelo menos 30 mil prédios/casas vazias.⁴⁴ O resultado desses dois fatores (falta de moradia e edifícios vazios) é a ocupação organizada recorrente. Trata-se de uma solução provisória que claramente incomoda o município.

Uma estratégia não oficial que foi adotada é a de criminalizar indiretamente os ocupantes, listando aqueles que cometeram crimes menores ou os seguindo para investigar e identificar qualquer comportamento ilícito. Se o reconhecimento facial for implementado, isso se tornará uma caça aos grupos vulneráveis que vivem nas cidades.

Essa tecnologia foi usada no edifício da Nelson Management Group, então há uma grande probabilidade de que a mesma tecnologia possa ser usada para proteger propriedades públicas e privadas de "intrusos" em edifícios vazios, mas também para aprimorar um Sistema de Reconhecimento Facial em guetos.

Além de aplicações em contextos específicos há falhas de natureza transversal, como o denominado racismo algorítmico, que evidencia uma situação "sociotécnica [...] concentradora, menos

⁴⁴ G1. Prédios abandonados e pessoas em busca de moradia: as contradições no Centro de SP. 09 de junho de 2021. Disponível em: <https://g1.globo.com/profissao-reporter/noticia/2021/06/09/predios-abandonados-e-pessoas-em-busca-de-moradia-as-contradicoes-no-centro-de-sp.ghtml>. Acesso em: 02 set. 2023.

inclusiva e mais discriminadora do que a existente no período anterior à popularização da Internet” (SILVA 2022, p. 11).

Esta espécie de racismo se mostra presente nas falhas de reconhecimento facial de pessoas pretas e pardas, com resultado de um treinamento dos algoritmos com bancos de imagens predominantemente compostos por imagens de pessoas brancas.

Como identificado por Joy Buolamwini, a falta de "diversidade nas bases de treinamento" de algoritmos é "um dos principais problemas". Isso prejudica o aprendizado de máquina durante o reconhecimento de padrões e a posterior aplicação desses aprendizados em conjuntos de dados diferentes daqueles inicialmente utilizados no treinamento. Essas falhas de identificação com grupos específicos evidenciam uma vulgarização dos erros que afeta a imagem, a identidade e a autoestima dos sujeitos (SILVA 2022, p. 76-77).

Os riscos desta falha podem levar a um encadeamento de situações problemáticas, como prisões por falhas na identificação, eventual prisão preventiva, permanência em cárcere e danos adicionais com reverberações nem sempre sanáveis. Mantido o sigilo do caso e a alteração de dados, não identificando a parte dos autos, pode-se citar processo que tramita na Justiça Estadual de São Paulo como paradigma para a compreensão dos danos que a identificação errônea pode causar.

Em 2002, determinado indivíduo, de nome fictício Jonas, foi erroneamente identificado por uma vítima através de um caderno impresso com fotografias de suspeitos. Este caderno é constituído pelas forças de segurança. A polícia, durante abordagem, identifica Jonas como o suspeito procurado.

Jonas tem a prisão decretada e permanece em cárcere durante alguns dias. Contrai tuberculose, que é atestada pelo hospital da carceragem. Após certo tempo tem a liberdade decretada. Todavia, passa a sofrer com danos da doença e perde parte considerável da capacidade respiratória.

Este busca advogado, ajuíza uma ação de reparação de danos morais. Em 2007 a sentença nega o pedido de indenização. Apresentado o recurso, a TJSP concede o pedido, um valor de 50 mil reais. Cumprindo com o dever recursal, a Procuradoria ataca a forma

de cálculo dos juros, indicando que contra a Administração Pública a taxa mensal deve ser de 0,5% (meio por cento) ao mês e não de 1% (um por cento). O Recurso Especial é encaminhado ao Superior Tribunal de Justiça (STJ) e os autos, em razão de repercussão geral, ficam sobrestados.

Eis que, em 2017 o Supremo Tribunal Federal (STF), no Recurso Extraordinário 870.947 (STF 2017, p. 2), decide pela impossibilidade jurídica da utilização do índice de remuneração da caderneta de poupança como critério de correção monetária e que caberia a aplicação do art. 1º-F, da Lei 9.494/97, que indicava o percentual de 0,5% a.m. (meio por cento ao mês).

A decisão cascadeia para os processos do STJ e os autos de Jonas são devolvidos para o TJSP. Jonas apresenta procede com o pedido de entrada na fila de precatórios em 2018. Deferido o pedido para que a Administração Pública pague a condenação, Jonas entra na fila dos precatórios, cujo tempo médio de espera, em 2018, era de 14 anos: ou seja, em 2018 receberam os valores devidos aqueles que entraram na fila em 2004.

Considerando este cenário, é provável que Jonas receba a indenização em 2032. Trinta anos após a falha de reconhecimento. Uma opção seria o acordo com o erário para o recebimento de Precatório via Requisição de Pequeno Valor (RPV), cujo pagamento ocorre em até 90 (noventa) dias da data de apresentação à entidade devedora.

O teto máximo para pagamento por meio de RPVs é definido por lei própria de cada ente federativo, levando em conta as diferentes capacidades econômicas. No caso do Estado de São Paulo, condenações de até 440,214851 Ufesp (o equivalente a R\$ 12.154,33 em 2020) (TJSP 2023).

Na eventualidade de falhas no reconhecimento pelo algoritmo, as derivações do erro podem trazer prejuízos consideráveis aos sujeitos. Embora haja uma ampla percepção de que a prisão é algo natural, um ingrediente quase "essencial do nosso senso comum", há grupos mais propensos a sofrer o encarceramento por fatores culturais e históricos que se veem amplificados pelos vieses de algoritmos não treinados com imagens suficientes de determinados grupos. (DAVIS *apud* SILVA 2022, p. 126).

Governança de Dados

A distinção entre gerir e governar reside no fato de que gerir diz respeito a realizar as coisas corretamente, enquanto governar se relaciona a assegurar que as coisas certas sejam feitas (WODZINSKI *et al.* 2015). Deste modo a Governança de Dados desempenha um papel fundamental ao aproveitar o conhecimento e a expertise das pessoas para estabelecer políticas, atribuir responsabilidades, criar glossários, gerenciar metadados.

Além de otimizar fluxos de trabalho de dados em trânsito, monitorar a linhagem de dados e, assim, capacitar a organização a aprimorar a qualidade, descoberta e compreensão desses dados. Isso garante a gestão adequada dos ativos de dados, conforme destacado por John Ladley (2012).

No nível municipal, na Cidade de São Paulo (2020) a Controladoria Geral do Município, por meio do Encarregado pela Proteção de Dados Pessoais, de acordo com as atribuições legais estabelecidas pela Instrução Normativa nº 01, de 21 de julho de 2022, com suas alterações introduzidas pela Instrução Normativa nº 01, de 13 de janeiro de 2023, disponibilizou *layouts* para o "Mapeamento de Dados Pessoais" e o "Relatório de Impacto à Proteção de Dados Pessoais", e tomou outras providências para estabelecer padrões na implementação de práticas relacionadas à proteção de dados pessoais em toda a Administração Pública do Município.

Em nível estadual, foi estabelecido o Comitê Gestor de Governança de Dados e Informações do Estado de São Paulo⁴⁵, que é responsável por estabelecer políticas para o tratamento e armazenamento de informações coletadas dos cidadãos paulistas. Isso visa garantir a qualidade, segurança e transparência dos dados, que são essenciais para o desenvolvimento de políticas públicas voltadas às necessidades da população. Esse órgão está em conformidade com

⁴⁵ GOVERNO DO ESTADO DE SÃO PAULO. Saiba mais sobre Governança. Disponível em: [https://spmaisdigital.sp.gov.br/wps/portal/spmaisdigital/dados/governanca!/ut/p/z0/04_Sj9CPyKssy0xPLMnMz0vMAfIjo8ziTQx8XD28DYwMDZ29nQ0Cw9yNLX0NzYxcfQz0g_ULsh0VAfQp0a4!/>](https://spmaisdigital.sp.gov.br/wps/portal/spmaisdigital/dados/governanca!/ut/p/z0/04_Sj9CPyKssy0xPLMnMz0vMAfIjo8ziTQx8XD28DYwMDZ29nQ0Cw9yNLX0NzYxcfQz0g_ULsh0VAfQp0a4!/) Acesso em: 08 de setembro de 2023.

várias leis, incluindo a Lei de Acesso à Informação, o Marco Civil da Internet e a Lei Geral de Proteção de Dados Pessoais. A Central de Dados do Estado de São Paulo (CDESP), criada pelo Decreto nº 64.790/2020, reúne todas as informações coletadas por instituições públicas do Estado, proporcionando informações valiosas e transparentes sobre a população paulista.

Ocorre que houve incidentes de vazamento de dados, conforme noticiado pela imprensa. O primeiro a se destacar ocorreu em dezembro de 2022, quando o G1⁴⁶ noticiou que um executivo seria suspeito de vazar os dados pessoais de 13 milhões de passageiros cadastrados no Bilhete Único de São Paulo, contrariando a versão inicial de um ataque *hacker*. A Polícia Civil investigou o vazamento e descobriu que o suspeito era um dos diretores de uma empresa de tecnologia que gerenciava essas informações.

A fraude teria ocorrido após a empresa perder uma licitação para continuar a administrar os dados da SPTrans, empresa responsável pelo sistema de ônibus da cidade. A investigação busca determinar se outros diretores e a empresa estavam envolvidos nesse vazamento ilegal de dados para ganho comercial.

A Polícia Civil realizou buscas em endereços relacionados ao caso e encontrou evidências, incluindo um computador e um telefone celular, que podem ajudar a esclarecer como os dados dos passageiros do Bilhete Único foram parar na *darkweb*. As informações pessoais dos passageiros, como nomes, endereços, documentos e senhas, foram expostas, mas os créditos do bilhete não foram afetados.

O segundo tratou do envio de informações sigilosas do sistema DETECTA para membros de facção criminosa.⁴⁷ A Polícia Civil de

⁴⁶ G1. GALVÃO, César. Executivo de empresa é o responsável por vazar dados de 13 milhões de usuários do Bilhete Único, diz polícia. G1, 31 de janeiro de 2023. Disponível em: <https://g1.globo.com/sp/sao-paulo/noticia/2023/01/31/executivo-de-empresa-e-o-responsavel-por-vazar-dados-de-13-milhoes-de-usuarios-do-bilhete-unico-diz-policia.ghtml>. Acesso em: 23 de setembro de 2023.

⁴⁷ G1. Polícia de SP investiga como facção teve acesso ao sistema de câmeras de monitoramento do estado. 24 de março de 2023. Disponível em: <https://g1.globo.com/jornal-nacional/noticia/2023/03/24/policia-de-sp-investiga-como-facciao-teve-acesso-ao-sistema-de-cameras-de-monitoramento-do-estado.ghtml>. Acesso em: 23 de setembro de 2023.

São Paulo ainda investiga como uma facção criminosa teve acesso ao sistema de câmeras de monitoramento do Estado, que é usado para fiscalizar o trânsito e auxiliar a polícia na identificação de criminosos. O sistema combina imagens de mais de 3 mil câmeras com bancos de dados de várias instituições, e emitiu um alerta quando a facção acessou informações de um veículo usado em ações de inteligência. As autoridades descobriram que a facção estava planejando sequestrar e assassinar várias autoridades, incluindo o senador Sergio Moro, a deputada federal Rosângela Moro e outros funcionários da segurança pública. A investigação aponta para o envolvimento interno no vazamento de informações para a facção.

O terceiro caso, de setembro de 2023, tratou do comprometimento de credenciais de vários órgãos públicos.⁴⁸ Um grupo de jovens, aparentemente liderado por um adolescente de 14 anos (identificado como Adolescente B), foi descoberto comercializando mais de 20 milhões de *logins* e senhas de órgãos como a Justiça e as principais polícias do Brasil. As credenciais incluem mais de 3.600 *logins* e senhas do Tribunal de Justiça de São Paulo, 1.500 da Polícia Militar de São Paulo, 500 da Polícia Federal, quase 150 do Exército e 89 do Ministério Público de São Paulo.

O Adolescente B alegou ser responsável pela criação de um programa de computador capaz de invadir qualquer *site*, facilitando o acesso aos sistemas dessas instituições. Ele explorava falhas de segurança em *sites* oficiais e usava servidores privados para ocultar suas atividades. O grupo operava em diversas partes do Brasil, com conexões em Jaciara (MT), Blumenau (SC) e Curitiba.

Apesar da prisão inicial de cinco *hackers* durante a Operação Lotter - Fraudador Digital, alguns foram liberados, mas o Adolescente A foi apreendido novamente e está sob custódia na Fundação Casa. A empresa Discord, utilizada pelo grupo, afirmou adotar uma política de

⁴⁸ G1. Polícia desmantela quadrilha com mais de 20 milhões de logins e senhas da polícia, Exército e Justiça; adolescente de 14 anos é apontado como líder. 24/09/2023. Disponível em: <https://g1.globo.com/fantastico/noticia/2023/09/24/policia-desmantela-quadrilha-com-mais-de-20-milhoes-de-logins-e-senhas-da-policia-exercito-e-justica-adolescente-de-14-anos-e-apontado-como-lider.gh.html>. Acesso em 25 de setembro de 2023.

tolerância zero para atividades ilegais e colabora com as autoridades quando informada sobre tais condutas.

A discussão passa então para o potencial risco que uma coleta massificada de dados pode gerar para a população. Considerando que pode haver o vazamento por má intenção, a disponibilização de informações confidenciais para terceiros ou, ainda amplo risco de comprometimento de credenciais, um sistema de reconhecimento facial em tempo real, ainda que com exclusões periódicas, representa riscos aos direitos de um total não quantificável de sujeitos, bem como facilitaria o rastreo de forças de segurança.

Na eventualidade do comprometimento destes dados seria possível: (i) rastrear as forças de segurança; (ii) utilizar o banco para treinamento de outros algoritmos; (iii) utilizar algoritmos existentes em imagens captadas por circuitos não controlados pela Prefeitura, de modo a criar um sistema paralelo de identificação. Qualquer uma das hipóteses colocam em risco oficiais, membros do três poderes e demais sujeitos que sejam potenciais alvos.

Considerações finais

Devido aos riscos dessa tecnologia, a melhor opção, neste momento, é proibir o reconhecimento facial ou outro tipo de reconhecimento biométrico em espaços públicos ou em espaços nos quais seu uso permitiria uma vigilância em massa; leis, regulamentos e estatutos que proíbam a vigilância em áreas públicas; afastar o uso do reconhecimento facial como única forma de acessar direitos (econômicos, trabalhistas, educacionais, de moradia e outros); reforçar a necessidade da responsabilidade e da transparência públicas.

Nota-se uma necessidade de que as entidades públicas e privadas atentem-se às aplicações de sistemas algorítmicos que possam representar ataques aos direitos humanos e que possam comprometer a própria administração pública e suas estratégias. Antes que o direito ao protesto seja comprometido e o exercício de uma autodeterminação se veja deteriorado, outros riscos estarão no caminho, cabendo uma decisão clara quanto à necessidade e

viabilidade de implementação do reconhecimento facial em espaços públicos.

Referências

- ANISTIA INTERNACIONAL. **Inside the nypd's surveillance machine**. Disponível em: <<https://banthescan.amnesty.org/nyc/>>. Acesso em: 27 set. 2023
- ANISTIA INTERNACIONAL. **BAN THE SCAN HYDERABAD: Do you want your face to be used to track you?** Disponível em: <<https://banthescan.amnesty.org/hyderabad/>>. Acesso em: 20 de set. 2023.
- BBC. BRYANT, N. (2022) **Guerra na Ucrânia: o histórico 'contra-ataque' dos EUA e aliados ocidentais contra o 'agitador' Putin**. BBC, 16 de março de 2022. Disponível em: <<https://www.bbc.com/portuguese/internacional-60647379>>. Acesso em: 01 de set. de 2023.
- BRANDÃO, R.; ARBIX, G. (2022) **Artificial Intelligence, Ethics and Public Policy—The Use of Facial Recognition Systems in Public Transport in the Largest Brazilian Cities**. *Journal of Service Science and Management* 15(5), DOI: 10.4236/jssm.2022.155032.
- LARA, W.; GOMES, P. (2021). **Pandemia empurrou cerca de 20 mil famílias para habitações precárias e cidade de SP ganhou 24 novas favelas, diz secretaria**. G1. São Paulo, 24 de junho de 2021. Disponível em: <<https://g1.globo.com/sp/sao-paulo/noticia/2021/06/24/pandemia-empurrou-cerca-de-55-mil-familias-para-habitacoes-precarias-e-cidade-de-sp-ganhou-150-novas-favelas-diz-secretaria.ghtml>>. Acesso em: 08 set. 2023.
- GARCIA, M. (2023) **A desobediência civil como defesa da Constituição**. *Revista Brasileira de Direito Constitucional* 2.
- HUNTINGTON, S.P. (1991) **Democracy's Third Wave**. *Journal of Democracy*. Disponível em: <https://www.ned.org/docs/Samuel-P-Huntington-Democracy-Third-Wave.pdf>. Acesso em: 01 set. 2023.
- MILLER, K. (2022). **The Movement to Decolonize AI: Centering Dignity Over Dependency**. 21 mar. 2022. Disponível em: Disponível em: <<https://hai.stanford.edu/news/movement-decolonize-ai-centering-dignity-over-dependency>>. Acesso em: 30 set. 2023.
- ORGANIZAÇÃO DAS NAÇÕES UNIDAS. **Declaração Universal dos Direitos Humanos, 1948**. Disponível em: <https://www.unicef.org> Acesso em: 8 ago 2022.
- SILVA, T. (2022) **Racismo algorítmico: inteligência artificial e discriminação nas redes digitais**. São Paulo: Sesc Edições.
- PREFEITURA DO MUNICÍPIO DE SÃO PAULO. **Diário Oficial do Município de São Paulo**. Ano 65 São Paulo, terça-feira, 5 de maio de 2020 Número 84. Disponível em:

<https://diariooficial.prefeitura.sp.gov.br/md_epubli_memoria_arquivo.php?JiTER_eiX5XsxhuW-WmzsRL67rU6jM4NhF499XCp_WGugoy1ZhgwqDvl6NDXTV9fNioYa186M2YfhfM42bAcQPQ>. Acesso em: 30 de set. 2023.

PREFEITURA DO MUNICÍPIO DE SÃO PAULO. **Proteção de Dados Pessoais e Privacidade**. Disponível em: <https://www.prefeitura.sp.gov.br/cidade/secretarias/controladoria_geral/a_cgm/index.php?p=332358>. Acesso em: 2 de set. 2023.

THE NEW YORK TIMES. Kashmir Hill. **Wrongfully Accused by an Algorithm**. 24 de junho de 2020. Disponível em: <<https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html>>. Acesso em: 01 set. 2023.

SUPREMO TRIBUNAL FEDERAL. **Recurso Extraordinário 870.947 Sergipe**. Relator: Ministro Luiz Fux. Data: 20 de setembro de 2017. Disponível em: <<https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=14080728>>. Acesso em: 15 ago. 2023.

TRIBUNAL DE JUSTIÇA DE SÃO PAULO. **Precatórios**. Disponível em: <https://www.tjsp.jus.br/Precatorios>. Acesso em 29 set. 2023.

WODZINSKI, M. *et al.* **Building an Impactful Data Governance Program One Step at a Time**. Disponível em: <<https://www.brighttalk.com/webcast/10477/142503/building-an-impactful-data-governance-program-onestep-at-a-time>>. Acesso em: 26 set. 2023

Autores

Rodrigo Brandão. Rodrigo Brandão é pesquisador do Cetic.br/Nic.br e coordenador do Observatório da Inovação e Competitividade (OIC), sediado no Instituto de Estudos Avançados (IEA) da Universidade de São Paulo (USP).

Thiago Gomes Marcilio

Mestre em Filosofia do Direito pela Pontifícia Universidade Católica de São Paulo (PUC-SP). Pós-graduando em Direito Digital e Inovação pela Procuradoria Geral do Estado de São Paulo. Pesquisador do C4AI-USP-IBM-FAPESP e do Ethics4AI do IDP-Brasília. Advogado.

PROFILING À BRASILEIRA: **entre riscos reais e imediatas expectativas**

Cristina Godoy Bernardo de Oliveira
Rafael Meira Silva

Introdução

A crescente conscientização sobre o impacto do uso da inteligência artificial em diferentes setores da vida tem gerado uma série de questionamentos tanto acerca da necessidade do estabelecimento de limites quanto da publicidade dos métodos empregados para a coleta dos dados imprescindíveis para o seu desenvolvimento e da forma como ocorre o seu processamento.

Atualmente, verifica-se que um dos principais usos da inteligência artificial está focado na análise e no processamento de dados, sendo o seu principal objetivo prever ou influenciar comportamentos.

Dessa forma, nota-se que os dados transformaram-se em *commodities* (SARTOR e LAGIOIA 2020, p. 36), havendo um mercado valioso de dados que permitem a otimização de vendas, por exemplo.

Outra possível aplicação pode ser identificada tanto na busca pelo aumento da eficiência da Administração Pública quanto na prevenção ou mesmo repressão às tentativas de fraudes, mais especificamente à corrupção, tendo em vista que, em regra, todos os processos e procedimentos são estruturados, o que torna possível, por meio da análise do (pseudo)discurso, a identificação de padrões de comportamento que se repetem pelos agentes públicos envolvidos (MEIRA 2018).

No presente estudo, sob o ponto de vista conceitual, será indicado como o *profiling*, como técnica associada à inteligência artificial, deve ser compreendido e, portanto, quais são os respectivos métodos para coletas de dados e os respectivos limites jurídicos.

Assim, seja na seara privada ou, especialmente, na pública, a aplicabilidade da inteligência artificial deve ser estudada, adaptada e regulada conforme o setor específico em que será utilizada. Vislumbra-se, portanto, a necessidade do estudo do *profiling* como meio de aproveitamento das informações já existentes com o escopo de prever comportamentos para melhorar a prestação de serviços tendo a pessoa como principal beneficiária, sempre cuidando para não torná-la vítima da mesma tecnologia.

Construção conceitual de *profiling*

A associação entre o *profiling* e a inteligência artificial tanto potencializou o uso do primeiro para diferentes situações como ampliou a sua dinamicidade tendo em vista a capacidade de consideração de um número virtualmente ilimitado de variáveis que podem ser incluídas na formação do *profile* desejado. Assim, o assunto *profiling* é estudado em diferentes contextos e, por isso, em caráter preliminar, é necessária a sua conceituação técnica.

Inicialmente, verificamos que, enquanto a expressão *profiling* é comumente traduzida como perfil na língua portuguesa, ela é derivada da palavra italiana *profilo*, que significa traçar uma linha no contorno de um objeto (SARTOR e LAGIOIA 2020, p. 36).

Ademais, percebe-se que, independentemente do contexto linguístico, a principal preocupação dos *stakeholders* é a automatização de decisões por meio do *profiling* gerado a partir do aprendizado de máquina ou *machine learning*.

No que se refere à diferença entre *profile* e *profiling*, é importante destacar que o primeiro é compreendido como sendo a descrição de alguém sobre os principais aspectos e fatos que caracterizam a pessoa (SCHIAFFINO e AMANDI 2009, p. 193).

Já o termo *profiling* consiste na técnica empregada para identificar e representar indivíduos ou grupos em determinadas

categorias, permitindo a elaboração de modelos preditivos e influenciando comportamentos.

Ademais, conforme a definição fornecida por Gary T. Marx, o *profiling* emprega uma lógica indutiva que permite encontrar conexões que viabilizam o aumento da probabilidade de se identificar infrações por meio de pesquisas randômicas (MARX e REICHMAN 1984, p. 429).

Além disso, devemos recordar o ensinamento de Roger Clarke, que afirma que o *profiling* significa o processo de criação e de uso do profile gerado. Nessa linha, ele caracteriza-se como sendo uma “técnica de vigilância de dados (*dataveillance*)” em que um grupo de características comuns a uma determinada classe de pessoas é inferida por meio de dados passados, sendo que os titulares dos dados pessoais são analisados para se verificar quais possuem um perfil semelhante que se adequa às características de determinada classe (CLARKE 1993, p. 405).

Outra definição foi elaborada por Hildebrandt, segundo o qual o *profiling* corresponde ao processo de “descoberta” de padrões no banco de dados que pode ser empregado para “identificar ou representar um ser humano ou um grupo”. Além disso, é destacada a relevância da predição, ou seja, da capacidade de antecipar eventos futuros com base nos dados obtidos no passado (HILDEBRANDT 2009, p. 289).

Em suma, verificamos que a construção do conceito de *profiling* sob o aspecto técnico foi intensificada em decorrência da digitalização dos bancos de dados privados e públicos nas últimas décadas associada ao aperfeiçoamento da infraestrutura tecnológica, o que possibilitou a ampliação do processamento dos dados e, conseqüentemente, a geração de novas informações sobre as pessoas baseada em associações identificadas a partir do aprendizado de máquinas.

Conceito jurídico de *profiling*

No campo jurídico, há a tendência de conectar o conceito de *profiling* à técnica de automação do processamento de dados pessoais, viabilizando a previsão de ações, comportamentos e

tendências de uma pessoa ou de um grupo (FERRARIS *et al.* 2013, p.32).

Nessa linha, Valeria Ferraris (2013, p.33) realiza uma proposta de definição de *profiling*: trata-se de uma técnica de processamento de dados pessoais ou não pessoais automatizada, visando-se a desenvolver modelos preditivos de conhecimento com base nos dados utilizados, permitindo a construção de perfis que poderão ser aplicados em tomada de decisões.

Assim, pode-se salientar que, conforme Valeria Ferraris, no conceito de *profiling*, há dois termos que precisam ser compreendidos:

1) construção de *profile*: processo de dados que permite a descoberta de padrões inesperados, permitindo a criação de *profile*;

2) aplicação de *profile*: é o processo de identificação de um indivíduo como integrante de uma categoria, sendo que essa identificação permite a automatização da tomada de decisões.

A leitura realizada sobre a proposta conceitual de Valeria Ferraris demonstra possíveis aplicações iniciadas a partir da coleta de dados: primeiro, o *profile* é construído a partir da associação de dados e, em seguida, há a categorização cujos requisitos autorizam a identificação automatizada no primeiro nível e, conseqüentemente, a tomada de decisões no segundo nível.

Já no âmbito da União Europeia, identificamos o artigo 4(4) GDPR:

«Definição de perfis» qualquer forma de tratamento automatizado de dados pessoais que consista em utilizar esses dados pessoais para avaliar certos aspectos pessoais de uma pessoa singular, nomeadamente para analisar ou prever aspectos relacionados com o seu desempenho profissional, a sua situação económica, saúde, preferências pessoais, interesses, fiabilidade, comportamento, localização ou deslocações.

Nesse sentido, tendo como perspectiva a intervenção humana, a definição de perfil (*profiling*) é realizada mediante o uso de qualquer automatização do processamento de dados direcionada à

identificação de novas informações sobre uma pessoa ou grupo de pessoas (OLIVEIRA e SILVA 2021a).

Ademais, ao se utilizar o termo “avaliar” no conceito de *profiling*, pode-se notar que há um julgamento acerca de determinados aspectos individuais (OLIVEIRA e SILVA 2021a) e, por isso, vislumbra-se uma primeira situação em que se torna necessário o estabelecimento de limites.

No entanto, no que se refere à automação de decisões, o uso do *profile* apenas para fins estatísticos ou para a classificação de clientes não se caracteriza como *profiling* conforme o Regulamento Geral sobre Proteção de Dados (GDPR). Além disso, é necessária a automatização das decisões decorrente da classificação dos clientes, podendo existir ou não participação humana no processo de automatização.

Por fim, deve-se recordar a Recomendação CM/Rec(2010)13 da União Europeia 1.e. sobre o termo *profiling*:

Significa uma técnica de processamento de dados automatizados que consiste na aplicação de uma definição de perfil individual para tomar decisões concernentes a preferências pessoais dela ou dele, comportamentos ou atitudes.

Em suma, compreende-se que o conceito jurídico de *profiling* deverá ser periodicamente revisitado em decorrência do desenvolvimento tecnológico para avaliar as novas situações envolvendo a possível exposição das pessoas ou grupos de pessoas. Com isso, a crescente facilidade trazida pela automatização das decisões fundamentada em aprendizado de máquina deverá ser sempre reavaliada diante dos desafios da sua regulação jurídica, mas, para isso, torna-se imprescindível conhecer as suas diferentes fases.

Fases do *profiling* baseado em IA

A prática tanto na seara privada em razão da análise de contratos privados (especialmente bancários) quanto de documentos públicos (notadamente relatórios correccionais) permitiu confirmar a

metodologia necessária para a aplicação do profiling com três fases específicas (OLIVEIRA e SILVA 2021b):

- 1) coleta de dados;
- 2) análise dos dados coletados por meio da automatização; e
- 3) estabelecimento de padrões e identificação de características presentes ou futuras dos titulares de dados pessoais.

Inicialmente, a primeira fase (coleta) apenas se torna viável se atendidos três requisitos: a) disponibilidade de grande volume de dados; b) consentimento (ou autorização) para a sua obtenção; e c) padronização dos dados coletados que permita a sua manipulação e posterior análise.

Em seguida, a segunda fase (análise) pressupõe, tendo em vista o grande volume de dados que será necessário para viabilizar o aprendizado de máquina, o suporte tecnológico e o domínio sobre o seu funcionamento no âmbito jurídico para viabilizar a elaboração das perguntas essenciais para a compreensão dos possíveis desdobramentos, que por sua vez, permitirão a revisão dos algoritmos que serão aplicados na fase seguinte.

Por fim, a terceira fase (padronização) torna viável o aprendizado de máquina a partir da disponibilidade do grande volume de dados associado (primeira fase) às perguntas essenciais (segunda fase).

Assim, identificamos uma espiral de novas informações geradas de maneira consistente a partir da interação automatizada sobre os dados coletados com diferentes possibilidades que serão ditadas pelo algoritmo aplicado:

1. os questionamentos sobre os dados geram respostas cujas análises constituídas por informações (nível 1) acarretam novos questionamentos;

2. os novos questionamentos podem associar: a) apenas as novas informações entre si, gerando outras informações (nível 2); ou b) as novas informações e os dados coletados na fase 1, gerando, consequentemente, outras informações até então ocultas (nível 3).

Em determinado momento, como exercício teórico, postula-se que as futuras interações entre os diferentes níveis de informações obtidas escaparão à capacidade de análise humana e, por isso, é imprescindível que novos estudos na área jurídica sejam

desenvolvidos com o foco na compreensão das possíveis consequências e responsabilidades.

Disciplina do *profiling* na LGPD & GDPR

Após a identificação conceitual do termo *profiling* e das respectivas fases metodológicas, é possível avançar para as considerações normativas referentes a sua relação com a Lei Geral de Proteção de Dados do Brasil (LGPD).

Embora o legislador tenha buscado a inspiração para a elaboração da LGPD no Regulamento Geral Europeu sobre Proteção de Dados (GDPR), notam-se algumas diferenças em relação à disciplina do *profiling*:

- a) na LGPD, não há um conceito de definição de perfil;
- b) a LGPD não é tão restritiva como a GDPR em relação ao *profiling*;
- c) a LGPD não possui um dispositivo semelhante ao artigo 22 da GDPR, que estabelece o direito à não sujeição a decisões, exclusivamente, automatizadas, inclusive no que se refere à definição de perfis quando gerar efeitos na esfera jurídica do titular de dados pessoais.

Ademais, conforme disposto no §2º do artigo 12 da LGPD, há a proteção de dados por inferência, pois, segundo mencionado dispositivo:

§2º Poderão ser igualmente considerados como dados pessoais, para os fins desta Lei, aqueles utilizados para formação do perfil comportamental de determinada pessoa natural, se identificada.

Outro aspecto legal que deve ser ressaltado encontra-se no *caput* do artigo 20 da LGPD, que faculta ao titular dos dados pessoais pedir a revisão de decisões automatizadas que afetem os interesses do titular, incluindo a definição de perfil.

Em outras palavras, a Lei Geral de Proteção de Dados também protege expressamente os dados pessoais por inferência, notadamente, os decorrentes de processos de tratamento de dados pessoais automatizados.

Casos práticos sobre a obtenção de *profiles*

Embora a inteligência artificial já seja estudada há décadas, o seu uso foi ampliado em razão do *profiling* em situações que envolvem as interações registradas por meio da Internet, especialmente como decorrência do comércio eletrônico e do *marketing* digital.

Inicialmente, temos o caso dos denominados sistemas Adaptativos. Nele, o *profile* é importante para o efeito adaptativo, ou seja, o sistema adapta-se conforme o comportamento dos usuários. Deve-se ressaltar que, para os sistemas adaptativos, os interesses dos usuários é a principal parte do *profile* (BRUSILOVSKY e MILLÁN 2007, p.9).

Nessa linha, o segundo caso refere-se aos agentes inteligentes, cujo objetivo é a humanização do atendimento eletrônico por meio dos famosos “*bots*”. Assim, o *profile* permite a assistência personalizada, pois o agente aprende a ajudar o usuário das seguintes maneiras:

- (1) observando e imitando o usuário;
- (2) recebendo *feedback* do usuário;
- (3) o usuário dizendo o que o agente deve fazer; e
- (4) aprendendo por meio de outros agentes inteligentes ao assistirem outros usuários (MAES 1994, p.77).

Ademais, outro motivo usual para justificar a obtenção de *profiles* acontece em razão dos sistemas tutoriais inteligentes: o *profile* do usuário permite que o processo de aprendizado seja adaptado conforme o seu conhecimento e a sua forma de aprender (GARCÍA *et al.* 2007, p. 796).

Já em relação às aplicações de e-commerce, o *profile* é empregado para personalizar ofertas e recomendações de produtos para os usuários conforme os seus interesses (ADOMAVICIUS e TUZHILIN 2001, p. 76). Nesse sentido, os usuários podem descobrir produtos que sejam relevantes ou que facilitem as suas atividades por meio da personalização da propaganda.

Em seguida, no que se refere ao mercado de trabalho, deve-se apontar o sistema de gestão do conhecimento. A partir dessa

estratégia, as habilidades de um empregado e a sua *performance* registrada em uma empresa permitem que os administradores ou os gestores de projetos saibam alocar os empregados em setores e atividades que serão executadas de forma mais eficiente pelo indivíduo. Dessa maneira, o *profile* permite que o empregado assuma uma posição que seja a melhor para ele e para a empresa (SURE *et al.* 2000, p. 02).

Por fim, deve ser registrado o sistema de recomendação, cujo *profile* possui as avaliações de produtos, como, por exemplo, livros, notícias, filmes, músicas etc., sendo que as notas fornecidas pelo usuário permitem com que sejam recomendados novos produtos com base em seus interesses (RESNICK e VARIAN 1997, p. 56).

Em síntese, os casos práticos indicados são exemplos recentes e, inclusive, de uso massificado (especialmente no caso do comércio eletrônico) e disso decorre a necessidade do aprofundamento do tema *profiling*. No entanto, as facilidades oferecidas pela inteligência artificial nas diversas esferas da vida do indivíduo como consumidor, trabalhador ou (principalmente) cidadão, não podem ofuscar a realidade: a alta exposição a riscos que podem ou não ser conhecidos previamente.

Riscos do *profiling*

Antes de elencar as principais situações que implicam em riscos à sociedade em decorrência do *profiling*, é importante salientar que o Brasil carece de profissionais que saibam efetivamente realizar avaliações de risco em praticamente todas as áreas do conhecimento. Como exemplo da lacuna existente nessa área, a União Europeia, como condição para possibilitar a avaliação de riscos, realiza estudos (portanto, ainda não concluídos) sobre como será possível identificar esses casos e, para isso, investe milhões de euros. No Brasil, vale lembrar, até o momento, não há definição sobre o orçamento, a autoridade responsável pela sua implementação e, inclusive, a própria regulação e a sua estratégia estão permeadas por dúvidas.

Assim, a consideração sobre os potenciais riscos do *profiling* é assunto que deverá ser cuidadosamente analisado nos próximos

anos tendo em vista o estado da arte referente à inteligência artificial que avança rapidamente.

O primeiro risco identificado refere-se à dominação dos vendedores. Nesse caso, por meio da segmentação, os indivíduos não possuem acesso a todas as informações necessárias para decidirem acerca de determinada compra. Dessa forma, o que o consumidor verá seja na loja eletrônica ou nos anúncios é determinado pelo vendedor, não o inverso.

Em seguida, temos o risco da discriminação, isto é, por meio de modelos preditivos pautados no *profiling*, é possível verificar decisões automatizadas discriminatórias, sendo difícil a realização da defesa por parte do indivíduo lesado, por exemplo, a discriminação de gênero para a concessão de crédito (v.g., caso Apple Card).

Depois, há o risco da perseguição política de certos grupos. Por meio do *profiling*, é possível vigiar as ações de indivíduos considerados perigosos para determinado partido político que está no poder, por exemplo. Esse risco pode impactar negativamente na democracia.

Após, existe o risco de exclusão de indivíduos no mercado de trabalho, pois, é possível que empresas utilizem-se de *profiling* para realizar contratações de empregados, sendo que os indivíduos excluídos não poderão se defender das decisões dos empregadores, pois não saberão se foi feito um *profiling*, como foi realizado o *profiling*, quais dados foram empregados para o treinamento de máquina etc..

Por fim, não pode ser esquecido o risco do domínio de grandes empresas de tecnologia, porque elas já possuem elevada quantidade de dados dos indivíduos e, por definirem os seus perfis, passam a adquirir um poder elevadíssimo tanto na política quanto na sociedade, sendo muitas vezes mais fortes do que Estados.

Tendo em vista a diversidade dos riscos do *profiling* nos principais setores da vida (política, economia, trabalho etc.) e a assimetria de forças entre os detentores dos dados e aqueles que os manipulam, torna-se necessária a presença do Estado seja para esclarecer os cidadãos sobre os seus direitos seja para resolver eventuais problemas (atuais ou futuros) decorrentes da rápida implementação das novas tecnologias.

Conclusão

A ideia de tornar as coisas inteligentes (*smart things*) é eliminar a dificuldade de escolher e decidir. No entanto, deve-se perguntar: os indivíduos querem parar de escolher, delegando para as máquinas esta tarefa? Além disso, qual é o limite para que não haja uma perda da liberdade? São diversos os questionamentos que decorrem da implementação do *profiling*.

Desde a construção inicial do conceito de *profiling* até a sua positivação, em especial pela LGPD e pela legislação da União Europeia, verifica-se que a sociedade dá os seus primeiros passos no que se refere ao exercício do uso da sua liberdade no meio digital: ceder ou não ceder os dados pessoais? O que será feito com eles? Quem os manipulará para melhorar a vida na sociedade? Quem será responsável em caso de efetivação dos riscos? Há respostas, mas todas são preliminares.

Existe muita expectativa de implementação rápida para que os frutos da inteligência artificial sejam colhidos o quanto antes, mas não podem ser ignorados os riscos óbvios e já sentidos seja no mercado de trabalho ou no que tange à privacidade. É preciso discutir o assunto e é necessário pensar soluções que considerem a verdadeira dinâmica do *profiling*, pois, ele somente é possível se houver os seguintes ingredientes conjugados: um elevado volume de dados disponíveis, a capacidade tecnológica de processamento, a inteligência para a análise jurídica e a identificação de padrões nas diversas áreas da vida que possibilitem a categorização das pessoas conforme as características comuns identificadas.

Na verdade, o *profiling* à brasileira apresenta-se como mais um futuro problema com o qual teremos que lidar. O cidadão sozinho ou mesmo na forma de sociedade organizada não terá como disputar espaços e gerar seus próprios relatórios para contestar as supostas verdades indicadas por grandes empresas (na maioria estrangeiras) que recebem diariamente os dados dos brasileiros quando indicam o que devemos ver na internet por exemplo.

A solução viável é a atuação consistente por parte do Estado na defesa dos interesses dos cidadãos para que o lema seja: *profiling*

com responsabilidade e, por isso, com observação integral da legislação brasileira, seja por empresas ou pela própria Administração Pública.

Em suma, o *profiling* possui um grande potencial para ser considerado benéfico, mas é preciso, com urgência, a atuação do Estado para estabelecer os limites necessários por meio da avaliação dos riscos gerados para que a sociedade possa, de fato, ser beneficiada pelas promessas de uma vida melhor amparada pelo desenvolvimento tecnológico.

Referências Bibliográficas

- ADOMAVICIUS, G., TUZHILIN, A. (2001) **Multidimensional Recommender Systems: A Data Warehousing Approach**. In: Fiege, L., Mühl, G., Wilhelm, U. (eds) Electronic Commerce. WELCOM 2001. Lecture Notes in Computer Science, vol. 2232. Springer, Berlin, Heidelberg. Disponível em: <https://doi.org/10.1007/3-540-45598-1_17>. Acesso em 10/09/2023.
- BRUSILOVSKY, P.; MILLÁN, E. (2007). **User Models for Adaptive Hypermedia and Adaptive Educational Systems**. In: P. Brusilovsky, A. Kobsa, & W. Nejdl (Eds.), The Adaptive Web (pp. 3-53). Berlin: Springer-Verlag. Disponível em: <http://dx.doi.org/10.1007/978-3-540-72079-9_1>. Acesso em 03/09/2023.
- CLARKE, R. A. (1993) **Profiling: A Hidden Challenge to the Regulation of Data Surveillance**. In: Journal of Law and Information Science 4,2. Disponível em: <<https://openresearch-repository.anu.edu.au/bitstream/1885/46248/31/07Paper06.pdf>>. Acesso em 01/09/2023.
- FERRARIS, V.; BOSCO, F.; CAFIERO, G.; D'ANGELO, E.; SULOYEVA, Y., **Defining Profiling** (December 11, 2013). Disponível em: <<http://dx.doi.org/10.2139/ssrn.2366564>>. Acesso em 30/08/2023.
- GARCÍA, Natasha. (2018). **The use of criminal profiling in cybercrime investigations**. Tese para: Master of Science in Cybersecurity.. Disponível em: <https://www.researchgate.net/publication/327187114_The_use_of_criminal_profiling_in_cybercrime_investigations>. Acesso em 10/09/2023.
- HILDEBRANDT, M. (2009). **Profiling and the Rule of Law** (January 23, 2009). In: Identity in Information Society (IDIS), Vol. 1, 2008. Disponível em: <<https://ssrn.com/abstract=1332076>>. Acesso em 01/09/2023.
- MAES, P. (1994). Agents that reduce work and information overload. *Communications of the ACM* 37(7):31–40. Disponível em: <<https://doi.org/10.1145/176789.176792>>. Acesso em 03/09/2023.

MARX, G.T.; REICHMAN, N.. Routinizing the Discovery of Secrets: Computers as Informants. *Sage Journals* 27(4). Disponível em: <<https://doi.org/10.1177/000276484027004003>>. Acesso em 10/09/2023.

MEIRA SILVA, R. (2018) **A relação da ética com a esfera pública: da *actio* ao discurso**. Tese (Doutorado) – Universidade de São Paulo, São Paulo, 2018. Di . Acesso em: 03 out. 2023.

OLIVEIRA, C.G.B.; SILVA, R.M. (2021a) **Inteligência Artificial e Proteção de Dados: Desafios e Debates - Parte 2**. Instituto Avançado de Proteção de Dados, Ribeirão Preto, 2021. Disponível em: <<https://iapd.org.br/inteligencia-artificial-e-protecao-de-dados-profiling/>>. Acesso em: 01 de setembro de 2023.

OLIVEIRA, C.G.B.; SILVA, R.M. (2021b) **Inteligência Artificial e proteção de dados: Definição de perfil e desafios**. In: Migalhas de Proteção de Dados. Disponível em: <<https://www.migalhas.com.br/coluna/migalhas-de-protecao-de-dados/339300/ia-e-protecao-de-dados-definicao-de-perfil-e-desafios>>. Acesso em 30/08/2023.

RESNICK, P.; VARIAN, H.R. (1997) Recommender systems. *Communications of the ACM* 40(3): 56–58. Disponível em: <<https://doi.org/10.1145/245108.245121>>. Acesso em 05/09/2023.

SARTOR, G.; LAGIOIA, F. (2020) **The impact of the General Data Protection Regulation (GDPR) on artificial intelligence**. Bruxelas: European Parliament. Disponível em: <<https://dx.doi.org/10.2861/293>>. Acesso em 10/09/2023.

SCHIAFFINO, S.; AMANDI, A. (2009) Intelligent User Profiling. *Artificial Intelligence*, 5640: 193-216. Disponível em: <http://dx.doi.org/10.1007/978-3-642-03226-4_11>. Acesso em 12/09/2023.

SURE, Y., MAEDCHE, A., STAAB, S. (2000) **Leveraging corporate skill knowledge - From ProPer to OntoProPer**. Proc. 3rd International Conf. on Practical Aspects of Knowledge Management, Basel, Switzerland. Disponível em: <https://www.researchgate.net/publication/225070423_Leveraging_Corporate_Skill_Knowledge_-_From_ProPer_to_OntoProPer>. Acesso em 10/09/2023.

VIGDOR, N. (2019) Apple Card Investigated after gender discrimination complaints. *The New York Times*, 10 de novembro de 2019. Disponível em: <<https://www.nytimes.com/2019/11/10/business/Apple-credit-card-investigation.html>>. Acesso em 30/09/2023.

Autores

Cristina Godoy Bernardo de Oliveira

Professora de Direito da USP. Academic Visitor da Faculty of Law of the University of Oxford. Pós-doutora pela Université Paris I Panthéon-Sorbonne. Líder do Grupo de Pesquisa Direito, Ética e IA da USP - CNPq. Coordenadora do Grupo de Pesquisa "Tech Law" do IEA/USP. Membro fundador do IAPD.

Rafael Meira Silva

Advogado. Graduado e Doutor em Filosofia do Direito pela Faculdade de Direito da USP, com estágio doutoral pela Université Paris 1 Panthéon-Sorbonne pelo programa de doutorado sanduíche pela CAPES (2014/2015). Membro fundador IAPD.

NEURODIREITOS: a tutela jurídica à mente humana

Camila Pintarelli

Introdução

O fascínio pela interação dos seres humanos com a robótica acompanha a sociedade desde há muito, bastando lembrar dos desenhos e maquinários elaborados por Leonardo da Vinci, no século XV, para constatar que a curiosidade humana sempre nos impeliu à busca de mecanismos capazes de implementar nossas vicissitudes por meio de instrumentos ou dispositivos exógenos, integrados, de alguma forma, à nossa capacidade pensante.

Na medicina, o prestígio às possibilidades da interação da robótica na vida humana é realçado há muitos anos, desde os simples exemplos do marcapasso e do *stent*, assim como a própria evolução dos estudos sobre a neurociência, que partiram, em 1878, da análise do cérebro de um animal por Richard Caton, e chegaram rapidamente, em 1929, na análise da eletroencefalograma humana por Hans Berger, tida como a primeira mensuração da atividade elétrica cerebral humana, precursora do chamado *brain-computer interface* (BCI).

Mais recentemente, o surgimento de técnicas de neuroimagem não invasivas provocou revolução na compreensão do cérebro humano (YUSTE e KAUSEL 2022, p.287), o que foi intensificado na atual quadra com a considerável ampliação do espectro de neurotecnologias de imagem disponíveis tanto no âmbito clínico como para acesso comercial, sendo exemplos a ressonância magnética (MRI), a tomografia axial computadorizada (CAT), a tomografia por emissão de pósitrons (PET) e a ressonância magnética funcional (fMRI).

O intenso desenvolvimento da neurotecnologia, com a evolução das técnicas de interface conectiva entre o cérebro e a máquina, está expandindo verdadeiramente as fronteiras da ação humana, gerando esperança e expectativa sobre seus impactos positivos na economia e, na área da saúde, no tratamento e na cura de doenças, sendo digno de nota o exemplo ainda vivo na memória de todos o episódio vivenciado na Copa do Mundo de 2014, quando um jovem paraplégico deu o chute inicial da abertura daquele Mundial, contando com a ajuda do exoesqueleto BRA-Santos Dumont 1, vestimenta robótica passível de ser controlada pela atividade cerebral do paciente, captada por meio de uma touca com sensores aplicados ao couro cabeludo.

São avanços que podem e devem ser celebrados, pois trazem esperança de superação das fragilidades do ser humano, ao mesmo tempo em que reafirmam a histórica e ínsita vocação humana em buscar, por meio do conhecimento, o aprimoramento de nossos pares, o que remonta a tempos de outrora, como vimos inicialmente.

Paralelamente aos inegáveis avanços à vida humana que o progresso neurotecnológico traz, é fato que a consolidação de tais tecnologias está permitindo que algumas de suas aplicações possam se ofertadas em escala comercial, o que traz à tona reflexões sobre as consequências éticas e sociais da disseminação de seu uso.

Em recente artigo, professores catedráticos da Universidade de Columbia (EUA), da Pontifícia Universidade Católica do Chile e do Centro de Pesquisas Psíquicas Donostia da Espanha externaram fundada preocupação com os limites éticos e jurídicos envolvidos no contexto da miríade de possibilidades de acesso, coleta, compartilhamento e manipulação de dados cerebrais (YUSTE e KAUSEL 2022, p.287), preocupação esta que vem encontrando eco mundial.

Acresce-se a isso o fato de que os métodos de análise de dados neurais, por acompanharem as demais inovações tecnológicas, estão se tornando cada vez mais complexos dada sua integração com ferramentas de inteligência artificial e *machine learning*, o que permite que algoritmos passem a ser usados para perfilar e analisar esse conjunto de informações que a neurotecnologia é capaz de extrair da mente humana.

Tais questões éticas podem, em uma primeira abordagem, ser analisadas e conduzidas a partir da introdução, no próprio desenvolvimento e aprimoramento de tais pesquisas, de balizas filosóficas aptas a viabilizar a criação de *standards* éticos voltados a orientar e iluminar fronteiras relacionadas à interação desses dispositivos com o ser humano e ao uso dos dados neurais por outros sistemas autômatos de tomada de decisão.

Em outro enfoque, essas mesmas questões podem, igualmente, ser tratadas por meio do estudo e da implementação de políticas regulatórias aptas a nortear seu uso e, mais do que isso, voltadas à proteção e à preservação dos direitos subjetivos envolvidos, em uma perspectiva que leva em consideração a interdisciplinaridade da neurotecnologia e seus naturais reflexos na ordem jurídica, especialmente na compreensão da dignidade da pessoa humana e dos direitos humanos.

Essa é justamente a abordagem aqui explorada, reputando ser fundamental que meditemos acerca da expansão da compreensão jurídica e normativa da dignidade humana nesse novo contexto digital, a fim de garantir que, ao se reconhecer o existente e inevitável dinamismo tecno-social com o sujeito humano, o desenvolvimento científico e tecnológico se dê a serviço da pessoa humana e com respeito à vida, à igualdade e à liberdade.

Trata-se de assunto de grande importância no debate jus acadêmico, revestido, ainda, de considerável ineditismo fático e doutrinário, que mostram não apenas a necessidade da abordagem ora pretendida, como a possibilidade de o Brasil – como um dos principais celeiros, mercados e destinatários globais dos progressos da tecnologia – assumir protagonismo global ao reafirmar o giro kantiano no contexto das evoluções tecnológicas e científicas.

Marcos históricos e potencialidades regulatórias

Como vimos inicialmente, a interação do homem com a robótica e com a tecnologia colore a história social desde há muito, e os recentes e mais grandiosos impulsos nesse assunto, responsáveis pelas grandes revoluções e aprimoramentos em prol da vida humana

que vêm ocorrendo, decorrem certamente do conhecimento cada vez maior do cérebro humano e do funcionamento da atividade neural.

Um dos principais marcos a esse respeito é o revolucionário projeto BRAIN Initiative – Brain Research Through Advancing Innovative Neurotechnologies. Em 02 de abril de 2013, os Estados Unidos, sob a gestão do então Presidente Barack Obama, anunciaram o início desse programa, cujo objetivo seria e é o de promover a pesquisa, o desenvolvimento e a aplicação de tecnologias inovadoras voltadas à criação de uma compreensão dinâmica do funcionamento do cérebro humano, projeto cuja magnitude e importância são constatadas pela composição do quadro de seus pesquisadores, que reúne notáveis e renomados profissionais das mais importantes universidades e centros de pesquisas do mundo.

A importância do programa foi acompanhada, logo após o seu lançamento, por preocupações acerca das consequências e dos limites éticos de suas pesquisas e resultados. O próprio Presidente Barack Obama externou tais inquietudes, a ponto de encaminhar questionamento à Comissão Presidencial para o Estudo de Questões Bioéticas dos Estados Unidos buscando orientações acerca de tais questões, na qual enfatizou os possíveis impactos que as descobertas envolvendo o cérebro humano poderiam trazer à liberdade e à igualdade humanas.

New technologies to better visualize the brain and understand how it works promise to speed the discovery of new ways to treat and prevent brain disorders, (...). Advances in neuroscience can also raise ethical and legal issues that require reflection and analysis. (...) we should consider the potential implications of the discoveries that we expect will flow from studies of the brain, and some of the questions that may be raised by those findings and their applications – questions, for example, relating to privacy, personal agency, and moral responsibility for one's actions; questions about stigmatization and discrimination based on neurological measures of intelligence or other traits; and questions about the appropriate use of neuroscience in the criminal-justice system, among others. It will also be important to consider these types of

questions as they relate to different life stages, from infancy through old age.^{49 50}

A resposta ao questionamento feito resultou no estudo *Gray Matters: Integrative Approaches for Neuroscience, Ethics, and Society*⁵¹, oferecendo quatro recomendações à Casa Branca e à BRAIN Initiative, quais sejam, integrar a ética de forma explícita e o quanto antes nas pesquisas, avaliar as abordagens existentes e que possam existir para a integração ética, integrar ética e ciência em todos os níveis de educação, e incluir explicitamente as perspectivas éticas nos órgãos de revisão e consultivos.

Em síntese, a Comissão concluiu que a ênfase neurocientífica deve ser sempre a busca pela integração humana, e não na intervenção⁵² no ser humano.

⁴⁹ Carta do Presidente Barack Obama à Comissão Presidencial para o Estudo de Questões Bioéticas dos Estados Unidos. Disponível em <https://repository.library.georgetown.edu/bitstream/handle/10822/709231/Gray%20Matters%20Vol%201.pdf?sequence=1>, acesso em 04 de agosto de 2022. Importante notar a preocupação, já nessa comunicação presidencial, com o viés discriminatório da tecnologia, bem como o impacto que ela pode causar à integridade mental nos diferentes estágios da vida humana, preocupação esta que norteou a edição da já citada lei californiana sobre o código de design de tecnologias virtuais a crianças.

⁵⁰ Nota editorial (tradução): As novas tecnologias que permitem visualizar melhor o cérebro e compreender o seu funcionamento prometem acelerar a descoberta de novas formas de tratamento e prevenção de doenças cerebrais, (...). Os avanços da neurociência podem também levantar questões éticas e jurídicas que exigem reflexão e análise. (...) devemos considerar as potenciais implicações das descobertas que esperamos que resultem dos estudos sobre o cérebro e algumas das questões que podem ser levantadas por essas descobertas e pelas suas aplicações - questões, por exemplo, relacionadas com a privacidade, o arbítrio pessoal e a responsabilidade moral pelas nossas ações; questões sobre a estigmatização e a discriminação com base em medidas neurológicas de inteligência ou outras características; e questões sobre a utilização adequada da neurociência no sistema de justiça penal, entre outras. Será também importante considerar este tipo de questões no que se refere às diferentes fases da vida, desde a infância até à velhice.

⁵¹ Disponível em <https://repository.library.georgetown.edu/bitstream/handle/10822/709231/Gray%20Matters%20Vol%201.pdf?sequence=1>, acesso em 04 de setembro de 2022.

⁵² Disponível em <https://bioethicsarchive.georgetown.edu/pcsbi/node/3581.html>, acesso em 04 de agosto de 2022.

Em 2017, países como a China, a Coreia do Sul, o Japão, o Canadá e a Austrália, assim como a União Europeia, se uniram na chamada International Brain Initiative⁵³, com propósitos semelhantes à BRAIN Initiative norte-americana, em reforço ao movimento mundial pelo salto qualitativo de interação da tecnologia ao ser humano via conhecimento das atividades cerebrais, o que está proporcionando grandes desenvolvimentos a tecnologias como a *Deep Brain Stimulation* (DBS), a *Transcranial Magnetic Stimulation* (TMS) e *Brain-Computer Interfaces* (BCIs), que passaram a apresentar resultados promissores para o tratamento de doenças neurológicas.

Ao mesmo tempo em que inúmeras possibilidades de cura e melhoramento da vida humana passam a ser contempladas com o avanço da neurotecnologia, conforme falamos inicialmente, fato é que o aprimoramento e a sedimentação do uso destas tecnologias, com a possibilidade cada vez maior de ampliação de sua oferta em escala comercial, tornam candente o impacto, por vezes pouco visível ou previsível, causado pelo desenvolvimento científico e tecnológico à vida em sociedade e ao sujeito humano (LÓPEZ-SILVA e VALERA 2022, p.25).

Para além dos exemplos acima transcritos, o uso comercial das neurotecnologias traz consigo outra gama de situações práticas. Na área de *marketing*, a neurotecnologia de imagem já é usada para coletar dados do cérebro a fim de conduzir os processos de decisão do consumidor, o que contribui para a elaboração de campanhas publicitárias. Ainda, no campo eleitoral, essas ferramentas estão sendo utilizadas para mensurar a popularidade implícita dos candidatos nos eleitores, o que assume importância diante da existência de pesquisas que indicam que a primeira impressão visual do candidato é deveras impactante, sendo pouco relevante o desenvolvimento da campanha e os discursos empreendidos posteriormente (LAVAZZA 2022, p.83).

As situações narradas apontam que a evolução neurocientífica, intensificada por sua integração com a inteligência artificial e com instrumentos de *machine learning*, está criando

⁵³ Disponível em <https://www.internationalbraininitiative.org/>, acesso em 04 de agosto de 2022.

diariamente novos espaços axiológicos, antropológicos (CUOZZO 2022, p.118-119), éticos e de ação humana, pois é a autodeterminação e o livre arbítrio humanos que estão em debate.

Nesse sentido, os desafios sem precedentes trazidos pela neurotecnologia podem e devem ser interpretados à vista dos direitos humanos, de forma a buscar a compreensão da efetiva tutela da liberdade, da igualdade e, acima disso, da dignidade da pessoa humana, inclusive para perquirir se os conceitos atualmente presentes nos marcos regulatórios internacionais são capazes de acomodar a realidade ora vivenciada.

Encontramos nuances do desenvolvimento científico e tecnológico resguardadas em alguns diplomas internacionais, cabendo o destaque, por exemplo, ao artigo 27, da Declaração Universal dos Direitos do Homem⁵⁴ – que prevê o direito do ser humano em participar do progresso científico e de seus benefícios –, bem assim ao artigo 15, do Pacto Internacional sobre Direitos Econômicos, Sociais e Culturais⁵⁵, que contempla semelhante previsão.

Outro marco jurídico internacional importante nesse contexto é a Declaração Universal sobre Bioética e Direitos Humanos, cujo artigo 8º traz o respeito à vulnerabilidade humana e à integridade pessoal.⁵⁶ A vulnerabilidade prevista no tratado pode ser interpretada como condição inerente à natureza humana, isto é, à fragilidade e à corporeidade que fazem com que os seres humanos necessariamente dependam, para seu aprimoramento ôntico, das relações sociais e do ambiente que integram (VALERA 2022, p.191-192).

⁵⁴ Artigo 27.1., Declaração Universal dos Direitos do Homem – Todo ser humano tem o direito de participar livremente da vida cultural da comunidade, de fruir as artes e de participar do progresso científico e de seus benefícios.

⁵⁵ Artigo 15.1.b., Pacto Internacional sobre Direitos Econômicos, Sociais e Culturais - Os Estados Partes do presente Pacto reconhecem a cada indivíduo o direito de: (...) b) Desfrutar o progresso científico e suas aplicações.

⁵⁶ Artigo 8º, Declaração Universal sobre Bioética e Direitos Humanos – Respeito pela Vulnerabilidade Humana e pela Integridade Individual: A vulnerabilidade humana deve ser levada em consideração na aplicação e no avanço do conhecimento científico, das práticas médicas e de tecnologias associadas. Indivíduos e grupos de vulnerabilidade específica devem ser protegidos e a integridade individual de cada um deve ser respeitada.

Em uma perspectiva filosófica, a vulnerabilidade é condição humana inerente à sua existência em sua finitude e fragilidade, de tal maneira que não pode ser superada ou eliminada. Ao se reconhecerem como vulneráveis, as pessoas compreendem a vulnerabilidade do outro, assim como a necessidade do cuidado, da responsabilidade e da solidariedade, e não a exploração dessa condição por outrem (MORAIS e MONTEIRO 2017, p.311-319).

Daí a relevância de se traçar algum paralelo entre a tutela prevista na Declaração Universal sobre a Bioética e Direitos Humanos com a situação ora vivenciada no cenário de franco desenvolvimento de tecnologias capazes de promover incrementos cognitivos para superar nossas vulnerabilidades individuais, uma vez que há o risco de impacto a uma característica que nos torna humanos, qual seja, a essência humana em sua própria vulnerabilidade.

Outro documento internacional de relevo é a Declaração sobre o Direito ao Desenvolvimento da Organização das Nações Unidas, de 1986, que alça o indivíduo a protagonista do progresso econômico, social, cultural e político, devendo participar e dele se beneficiar, além de promover e proteger uma ordem política, social e econômica que seja igualmente favorável ao desenvolvimento humano de forma integral.⁵⁷

Os princípios compreendidos na Declaração de 1986 abarcam noções de proteção que transcendem o momento histórico de sua edição e alcançam a realidade tecnológica dos dias atuais, notadamente a igualdade e a inclusão no contexto do desenvolvimento, o que assume especial importância diante da possibilidade de oferta de meios de incremento cognitivo via

⁵⁷ Article 2 - 1. The human person is the central subject of development and should be the active participant and beneficiary of the right to development. 2. All human beings have a responsibility for development, individually and collectively, taking into account the need for full respect for their human rights and fundamental freedoms as well as their duties to the community, which alone can ensure the free and complete fulfilment of the human being, and they should therefore promote and protect an appropriate political, social and economic order for development. Declaração sobre o Direito ao Desenvolvimento da Organização das Nações Unidas, aprovada pela Resolução 41/128, durante a 97ª reunião plenária, aos 04 de dezembro de 1986. Disponível em <<http://www.un.org/documents/ga/res/41/a41r128.htm>>, acesso aos 26 de julho de 2016.

neurotecnologia. Outro ponto de grande destaque é a transparência e *accountability*, necessárias para afiançar a confiabilidade das relações sociais desenvolvidas em meio às novas tecnologias.

A despeito de tais referenciais, pesquisas indicam que os tratados internacionais existentes correlatos a este assunto não apresentam tutela jurídica suficiente para proteger a integridade física e mental diante do desenvolvimento da neurotecnologia. É o que concluiu recente análise da NeuroRights Foundation, ligada à Universidade de Columbia, nos Estados Unidos, uma das precursoras globais em estudos desse matiz:

(...) existing international human rights treaties are currently unprepared to protect neurorights. Nevertheless, as described in detail in our findings, rapid advances in neurotechnology are no longer science fiction – they are science. It is urgent that the UN play a leading role globally to embrace these exciting innovations while protecting human rights and ensuring the ethical development of neurotechnology. (Genser *et al.* 2022)⁵⁸

Referida pesquisa reconhece que os hiatos de proteção no cenário do direito internacional dos direitos humanos em relação à neurotecnologia podem ser endereçados valendo-se de expressões ou mesmo ideias contidas em alguns tratados já existentes, tal como sugerimos anteriormente. É o caso da própria Declaração Universal sobre Bioética e Direitos Humanos e da Declaração Internacional sobre Dados Genéticos Humanos, que podem atuar como diretrizes para a coleta e o armazenamento de dados cerebrais, bem assim indicar linhas bases em relação ao consentimento humano envolvido em tais procedimentos.

Contudo, a conclusão alcançada indica a necessidade premente de que novas terminologias de linguagem sejam

⁵⁸ Nota editorial (tradução): Os tratados internacionais de direitos humanos existentes não estão atualmente preparados para proteger os neurodireitos. No entanto, tal como descrito em pormenor nas nossas conclusões, os rápidos avanços na neurotecnologia já não são ficção científica - são ciência. É urgente que as Nações Unidas desempenhem um papel de liderança a nível mundial para acolher estas inovações empolgantes, protegendo simultaneamente os direitos humanos e assegurando o desenvolvimento ético da neurotecnologia.

incorporadas aos tratados internacionais de modo a contemplar uma proteção jurídica mais específica à gama de situações práticas que emergem do uso das diferentes tecnologias neurais, especialmente aquelas que envolvem a transferência de dados cerebrais, em nítida demonstração de que os conceitos atuais de liberdade e igualdade precisam ter sua compreensão ampliada de modo a resguardar a dignidade da pessoa humana.

A busca por novos direitos e pelo aprimoramento dos tratados internacionais que temos é um movimento natural e cíclico das sociedades, cabendo destacar que o reconhecimento de que as alterações sociais exigem o perfilamento de novos direitos é tido pela melhor doutrina jurídica como um dos mais vigorosos sinais de amadurecimento social (ZAGREBELSKY 2007, p.84).

Na ausência de instrumentos internacionais específicos, as manifestações globais regulatórias sobre o tema, apesar de serem extremamente recentes, reforçam essa sinalização de amadurecimento social e de busca por novos direitos, indicando, ainda, a consciência coletiva dos atores internacionais sobre a necessidade de conferir balizas orientativas, cogentes ou não, a respeito do assunto.

Nesse sentido, a Organização para a Cooperação e Desenvolvimento Econômico – OCDE pode ser considerada pioneira no cenário das organizações intergovernamentais ao emitir, em dezembro de 2019, a “Recomendação sobre Inovação Responsável em Neurotecnologia”⁵⁹, tida como o primeiro marco regulatório internacional sobre a matéria. A citada Recomendação reconhece a necessidade de padrões normativos internacionais para o assunto e foi construída em torno de nove princípios, visando a, em última análise, guiar os Estados-nação e pesquisadores de modo a antecipar e a endereçar os desafios éticos, legais e sociais trazidos pelo desenvolvimento da neurotecnologia.

Nessa mesma vereda, em outubro de 2020, o Parlamento Europeu aprovou resolução com recomendações sobre os aspectos éticos da inteligência artificial, da robótica e de tecnologias conexas,

⁵⁹ Disponível em <https://www.oecd.org/science/recommendation-on-responsible-innovation-in-neurotechnology.htm>, acesso em 20 de agosto de 2022.

buscando, em linhas gerais, reforçar a confiança social e exigir transparência e melhores fluxos de informação entre os cidadãos e as organizações e empresas que desenvolvem, implantam e utilizam tais tecnologias,⁶⁰ o que pode ser considerado um grande salto regulatório global para a matéria.

A Espanha adotou, em julho de 2021, Carta de Direitos Digitais, cujo texto contém uma série de princípios e direitos para guiar o futuro regulatório de políticas públicas, de forma a garantir proteção individual e coletiva nos novos cenários digitais. Dentre os pontos contidos em referida Carta, destacam-se a previsão de direitos digitais no uso de neurotecnologias e, também, a importância da tutela à privacidade e à autodeterminação mentais, bem como a proteção contra a discriminação.⁶¹

Na América Latina, o Chile é, até agora, o único país que avançou na análise do tema, expressando preocupação com a necessidade de advento de marco regulatório capaz de reconhecer novos direitos humanos à vista do desenvolvimento tecnológico atualmente experimentado.

Ao aprovar a Lei nº 21.383, publicada em 25 de outubro de 2021, o parlamento chileno buscou alterar a Constituição do país para contemplar o direito à neuroproteção como decorrência da proteção à dignidade humana, resguardando o direito à privacidade mental, à identidade e à autonomia pessoal, ao livre arbítrio e à autodeterminação, ao acesso equitativo a mecanismos de incremento cognitivo cerebral, e à proteção contra o preconceito algorítmico em processos automatizados de tomada de decisão. Essas previsões protetivas do legislativo chileno vêm sendo classificadas conceitualmente como neurodireitos.

Em março de 2023, o Comitê Jurídico Interamericano, da Organização dos Estados Americanos, deu novo passo no tema e aprovou a Declaração de Princípios Interamericanos em Matéria de Neurociências, Neurotecnologias e Direitos Humanos.

⁶⁰ Disponível em https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_PT.html, acesso em 20 de agosto de 2022.

⁶¹ Disponível em https://www.lamoncloa.gob.es/presidente/actividades/Documents/2021/140721-Carta_Derechos_Digitales_RedEs.pdf, acesso em 10 de agosto de 2022.

Trazendo essas inquietações e esses movimentos normativos ao cenário jurídico brasileiro, é possível visualizarmos maneiras para integrar a proteção jurídica à mente humana em nosso ordenamento, de forma a reforçar e garantir outros direitos fundamentais correlatos, como a igualdade e a liberdade, cuja releitura à luz do contexto tecno-social demanda aperfeiçoamento constitucional capaz de atualizar nossa Lei Maior a essa nova realidade.

Nesse sentido, é oportuno lembrar que a ordem constitucional brasileira demonstra opção do constituinte de atrair à conformação conceitual normativa da dignidade humana o conceito de autodeterminação do ser humano. Com efeito, a dignidade humana na Constituição de 1988 é a conjugação de liberdade como um princípio de sociabilidade: ela afirma a capacidade humana de reger o próprio destino, expressando sua singularidade individual, ao mesmo tempo em que nega o isolamento, pois afirma também o enraizamento social do homem, posto que sua dignidade repousa na pluralidade e no seu agir conjunto (FERRAZ 1997, p.13-27).

Trata-se de escolha constitucional que realça, igualmente, o quanto dito antes acerca da presença da vulnerabilidade na construção do ser humano e, no caso, da dignidade humana, na medida em que demonstra a importância da interação plural do ser em sociedade para a afirmação do sujeito enquanto titular de sua própria infinitude.

Nossa Constituição abraça, assim, a ideia de autodeterminação humana como elemento conformador da própria dignidade do ser, o que assume especial importância no contexto do desenvolvimento das neurotecnologias e da preocupação com a criação de *standards* regulatórios aptos a orientar seu uso de modo a preservar o ser humano.

Em outras palavras, a proteção jurídica da mente, no cenário normativo constitucional brasileiro, passa a ser decorrência natural da própria opção conceitual de dignidade humana esposada pela Constituição da República que, apesar de editada em um momento histórico no qual sequer se previa o progresso científico a tais patamares, é capaz de acomodar modificações de tal ordem, sem que haja prejuízo às noções fundamentais que guarnecem os direitos nela previstos.

Um passo bastante relevante que foi dado na busca por incorporar, na Constituição da República, a proteção ao sujeito humano em meio ao contexto digital foi a Emenda Constitucional nº 17, de 2019, concebida no Senado Federal, que incluiu o inciso LXXIX no rol do artigo 5º, a fim de prever a proteção de dados pessoais dentre os direitos e garantias fundamentais.

Vemos, contudo, que o caminho de desenvolvimento das neurotecnologias está além da proteção de dados pessoais, alcançando propriamente a mente humana, na medida em que já se mostra possível o processamento e até mesmo o armazenamento de dados neurais. Não se está a falar apenas de dados pessoais, mas sim da própria integridade mental enquanto formadora da essência do ser humano, o que atinge a liberdade, a igualdade, a autodeterminação pessoal e a dignidade da pessoa humana.

Considerando a abertura conceitual constitucional conferida à dignidade da pessoa humana e a compreensão parlamentar pela necessidade de avanços na previsão de direitos e garantias fundamentais, conforme sinalização conferida na Emenda Constitucional nº 17, de 2019, parece-nos possível que o rol de direitos e garantias fundamentais seja novamente objeto de aprimoramentos, expressando o amadurecimento social brasileiro pelo reconhecimento de novos espaços axiológicos de proteção jurídica.

Não olvidamos que a sociedade brasileira ainda carece da concretude e efetividade de inúmeros direitos fundamentais já previstos na nossa Constituição e que a ausência de políticas públicas aptas a implementá-los torna a situação bastante preocupante em termos de desenvolvimento social e de desenvolvimento enquanto nação. Contudo, entendemos que não podemos nos prender a esse argumento para deixar de abordar a necessidade regulatória de um tema cada vez mais palpável e sensível em nossa realidade, cujos contornos talvez fiquem apenas mais visíveis quando já não for possível debater suas consequências éticas e normativas em relação à tutela da pessoa humana.

Assim, entendemos viável ao constituinte derivado que promova nova alteração ao artigo 5º, da Constituição Federal, de modo a trazer para o corpo constitucional brasileiro previsão de tutela

específica à mente humana no contexto das neurotecnologias, de modo a assegurar que o desenvolvimento científico e tecnológico garanta a integridade mental e a transparência algorítmica.

O Brasil, ao buscar integrar em sua Constituição norma de proteção específica à proteção jurídica da mente, se juntaria a outros atores globais e se colocaria em posição de protagonismo global no assunto ao demonstrar a exata compreensão do momento tecnológico vivenciado no mundo, mostrando responsabilidade com a pessoa humana ao mesmo tempo em que reafirmaria o compromisso com o desenvolvimento científico e tecnológico, em clara aderência a recomendações já emanadas de órgãos como a Organização para a Cooperação e Desenvolvimento Econômico – OCDE e a Organização dos Estados Americanos.

Conclusão

Vimos que o intenso desenvolvimento da neurotecnologia, com a evolução das técnicas de interface conectiva entre o cérebro e a máquina, está expandindo as fronteiras da ação humana, gerando esperança e expectativa sobre seus impactos positivos na economia e, na área da saúde, no tratamento e na cura de doenças.

Embora sejam avanços que devam ser celebrados, não podemos perder de vista os limites éticos envolvidos em tais procedimentos, pois trata-se de escalada evolutiva que diz respeito à própria integridade psíquica do ser humano, no que inserimos, também, as preocupações relativas à sedimentação dos processos automatizados de tomada de decisão via interação algorítmica, que também impactam a mente humana e demandam tutela jurídica específica.

A despeito da conexão temática, os atuais diplomas normativos internacionais não conferem proteção jurídica suficiente à conjuntura tecno-social vivenciada, existindo, todavia, pesquisas e iniciativas intergovernamentais, supranacionais e nacionais – como é o caso do Chile e da Espanha – contemplando balizas para a tutela da privacidade mental no contexto da disseminação do uso da neurotecnologia em escala comercial. Trata-se de preocupação que também atinge o cenário brasileiro.

Nesse sentido, considerando a conformação conceitual normativa conferida à dignidade da pessoa humana por nossa Constituição Federal e diante da vereda trilhada por outros atores globais, entendemos possível que o constituinte derivado tenha a sensibilidade em promover a expansão da compreensão jurídica da dignidade humana nesse novo contexto digital, incorporando à Constituição Federal brasileira semelhante proteção, a fim de complementar previsão já existente em relação aos dados pessoais e garantir que, diante do existente dinamismo tecno-social, o desenvolvimento científico e tecnológico se dê a serviço da pessoa humana e com respeito à vida, à igualdade e à liberdade.

Referências bibliográficas

- BOBBIO, N. (2004) **A Era dos Direitos**. Rio de Janeiro: Elsevier.
- CHAUDHARY, U., VLACHOS, I., ZIMMERMANN, J.B. et al. (2022) Spelling interface using intracortical signals in a completely locked-in patient enabled via auditory neurofeedback training. *Nat Commun*. 2022. Disponível em <https://doi.org/10.1038/s41467-022-28859-8>, acesso em 18 de agosto de 2022.
- CUOZZO, G. (2022) Responsibility: A Theory of Action Between Care for the World, Ethology and Art. In: LÓPEZ-SILVA, P. e VALERA, L. Protecting the Mind: Challenges in Law, Neuroprotection, and Neurorights. Switzerland : Springer.
- FERRAZ, T.S. (Jr.) (1997) Direito e Cidadania na Constituição Federal. *Revista da Procuradoria Geral do Estado de São Paulo* 47/48.
- GENSER, J.; HERRMANN, S.; YUSTE, R. (2022) International Human Rights Protection Gaps in the Age of Neurotechnology. Disponível em: <https://static1.squarespace.com/static/60e5c0c4c4f37276f4d458cf/t/6275130256dd5e2e11d4bd1b/1651839747023/Neurorights+Foundation+PUBLIC+Analysis+5.6.22.pdf>, acesso em 20 de agosto de 2022.
- LAVAZZA, A. (2022) Free Will and Autonomy in the Age of Neurotechnologies. In: LÓPEZ-SILVA, P. e VALERA, L., Protecting the Mind: Challenges in Law, Neuroprotection, and Neurorights. Switzerland: Springer.
- LÓPEZ-SILVA, P. (2022) The Concept of Mind in Neuroprotection Debate. In: LÓPEZ-SILVA, P. e VALERA, L. Protecting the Mind: Challenges in Law, Neuroprotection, and Neurorights. Switzerland : Springer.
- MORAIS, Talita C. A. de. MOTEIRO, Pedro Sadi. Conceitos de vulnerabilidade humana e integridade individual para a bioética. *Revista Bioética* 25(2): 311-

319, 2017, DOI: 10.1590/1983-80422017252191, acesso em 04 de agosto de 2022.

PIOVESAN, F. (2010) Direito ao Desenvolvimento – Desafios Contemporâneos. In: PIOVESAN, F. e SOARES, I.V.P. Direito ao Desenvolvimento. Belo Horizonte: Fórum.

SOMMAGGIO, P. (2022) Neuroscience, Neurolaw and Neurorights. In LÓPEZ-SILVA, P. e VALERA, L. Protecting the Mind: Challenges in Law, Neuroprotection, and Neurorights. Switzerland : Springer.

SU, C. ZHOU, H. GONG, L. TENG, B. GENG, F. HU, Y. (2021) Viewing personalized video clips recommended by TikTok activates default mode network and ventral tegmental area. *NeuroImage* 237: 118-136. Disponível em: <https://doi.org/10.1016/j.neuroimage.2021.118136>, acesso em 11 de agosto de 2022.

VALERA, L. (2022) Mental Integrity, Vulnerability and Brain Manipulations: A Bioethical Perspective. In: LÓPEZ-SILVA, P. e VALERA, L. Protecting the Mind: Challenges in Law, Neuroprotection, and Neurorights. Switzerland: Springer.

YUSTE, R.; ÁLAMOS, M.F.; KAUSEL, L. *et al.* (2022) A Technocratic Oath. In: LÓPEZ-SILVA, P. e VALERA, L. Protecting the Mind: Challenges in Law, Neuroprotection, and Neurorights. Switzerland: Springer.

ZAGREBELSKY, G. (2007) **El derecho dúctil**. 7.ed. Madrid: Trotta.

Autora

Camila Pintarelli

Doutora e Mestre em Direito (PUC/SP). Pós-Doutoranda em Direito (Universidade de Coimbra). Pesquisadora do Ethics4AI (Mackenzie/IDP). Procuradora do Estado de São Paulo.

A INTELIGÊNCIA ARTIFICIAL E O FUTURO DO TRABALHO E DA PROTEÇÃO SOCIAL

Alvaro A Comin

Introdução

Neste artigo abordamos dois temas que vêm dominando crescentemente a agenda de debates sobre o “futuro trabalho”: o impacto da Inteligência Artificial sobre os empregos e as profissões; e a expansão das plataformas de trabalho. Ambos os fenômenos estão ligados a uma mesma matriz de inovações tecnológicas e têm potencial para produzir grandes impactos sobre os mercados de trabalho, com consequências preocupantes para o futuro dos sistemas de proteção social construídos em torno do paradigma do trabalho assalariado regular.

Sabemos que as inovações tecnológicas tanto criam quanto destroem ocupações, mas nunca é possível saber, de antemão, qual será o saldo, até porque o volume em que os empregos serão criados não é determinado pela tecnologia apenas, mas pelo volume dos investimentos, do consumo, do comércio internacional e do crescimento econômico. Por outro lado, as mudanças tecnológicas afetam muito diretamente o *mix* de ocupações e, assim, também os tipos de conhecimentos e habilidades profissionais demandados, provocando ao mesmo tempo desvalorização e obsolescência de parte da força de trabalho e gargalos de formação e experiência profissional, que hipervalorizam qualificações escassas.

Para complicar mais as coisas, pelo caráter imaterial das novas atividades que mais se expandem e pelo alcance global das infraestruturas de comunicação e das corporações que as controlam, a correlação espacial entre oferta e demanda por trabalho tende a se

dissolver cada vez mais, de modo que os investimentos realizados em um determinado país ou região podem gerar empregos, principalmente, em outros países ou regiões.

Considerando o papel central que os mercados de trabalho desempenham na distribuição da renda e do consumo e na reprodução social de maneira ampliada, questões bastante cruciais sobre o futuro são incontornáveis. Em um horizonte de 20 ou 30 anos, haverá emprego para todas as pessoas que necessitam trabalhar para viver? A força de trabalho existente estará preparada para desempenhar as novas funções resultantes das inovações tecnológicas e nos modelos de negócios? Quais tipos de relações de trabalho predominarão?

As plataformas de trabalho, por sua vez, são elementos importantes nesta dinâmica de criação, destruição e realocização do trabalho, pois fornecem os nexos que permitem o aprofundamento da divisão do trabalho, articulando diferentes (novas e velhas) formas de arregimentação, controle e remuneração da força de trabalho, tornando-as mais flexíveis e mais desterritorializadas, maximizando os interesses das empresas e impulsionando uma série de novos modelos de negócios. A literatura sobre o assunto é praticamente unânime em apontar a associação entre as plataformas de trabalho e a tendência à diluição do emprego assalariado regular e formal em modalidades mais precárias de contratação onde estão ausentes os direitos tipicamente garantidos aos empregados pelas legislações trabalhistas (HUWS , 2017). No pior dos cenários, no futuro que nos aguarda haverá menos oportunidades de trabalho e elas serão mais precárias.

Na próxima seção deste artigo, discutiremos como a presente onda de inovações tecnológicas pode influenciar o volume e o perfil da demanda por trabalho. Na sequência, procuramos identificar como essas tendências se associam às formas de exploração do trabalho.

Automação e Inteligência Artificial: existe um futuro para o trabalho?

Historicamente, as tecnologias de automação substituíram sobretudo os trabalhos de natureza manual e rotineira, notadamente na indústria e na agricultura. E assim, o amplo e heterogêneo setor de serviços se consolida como o grande estuário das oportunidades de

trabalho, bem como das inovações voltadas para a automação de funções que reduzam a necessidade de trabalho humano, um objetivo permanente das empresas.

O desenvolvimento da Inteligência Artificial (AI) não é o único, mas é provavelmente o principal elemento tecnológico a influenciar o ritmo e a extensão da automação dos serviços e é, portanto, fator-chave no debate sobre o “futuro do trabalho”. Sistemas de IA são capazes de realizar operações que envolvem raciocínio lógico, tomada de decisões e resolução de problemas. Sua combinação com o poder de processamento dos computadores, o qual cresceu exponencialmente nas últimas décadas, somados às infraestruturas de telecomunicações e às redes digitais, fazem dela uma tecnologia de aplicação geral, que serve potencialmente a todas as demais atividades humanas, com um poder disruptivo que tem sido comparado ao da eletricidade no início do século XX (FORD 2015).

Na prática, a IA vem permitindo o avanço da automação em todos os setores da economia, da agricultura aos serviços de saúde, da indústria do petróleo aos serviços financeiros, da condução de veículos ao entretenimento e muito além (LANE e SAINT-MARTIN 2021). E, o que é o mais crucial, a IA vem avançando em área de atividades profissionais que até recentemente supúnhamos ser redutos invioláveis do trabalho humano, atividades que envolvem experimentação, pesquisa, tomada de decisões e até mesmo criatividade, como a redação de artigos de notícias e composição de sinfonias para orquestras (BRYNJOLFSSON e MCAFEE, 2011). Todas estas são atividades desempenhadas por trabalhadores com educação formal, qualificação e experiência profissionais elevadas, que normalmente desfrutam de bons contratos de trabalho e salários acima da média (AUTOR 2020).

Ainda mais expostos ao risco de desemprego tecnológico estão as ocupações que envolvem atividades de natureza técnica mais rotineiras e previsíveis, como arquivistas e almoxarifes, caixas e cobradores, vendedores, analistas de crédito, auxiliares de escritório e desenhistas, para ficar em apenas alguns exemplos, todas ocupações tipicamente desempenhadas em empresas e instituições, sob contratos regulares e estáveis de trabalho.

Em contraste com o que acontece com muitas ocupações que envolvem conhecimentos abstratos e raciocínio, que podem ser automatizados com recursos computacionais relativamente leves, ocupações que exigem esforços físicos não previsíveis e empenho das funções sensoriais e motoras, características, por exemplo, da construção civil, da agricultura de frutas e hortaliças e dos serviços domésticos, são muito mais difíceis de automatizar e demandam um poder computacional “muito mais pesado” (Autor, 2014). Via de regra, essas são também ocupações associadas a baixos salários e relações de trabalho informais, muito frequentemente desempenhadas por imigrantes nos países mais ricos ou pelos estratos mais pobres e menos escolarizados da força de trabalho nacional em países de renda média ou baixa. As dificuldades técnicas e as barreiras de custos envolvidas na automação destes tipos de atividades contrastam também com o avanço contínuo da automação na manufatura, onde o trabalho, igualmente de natureza manual, é predominantemente de tipo regular e formal.

Em conjunto, essas tendências indicam que o impacto das tecnologias de automação tem sido mais intenso sobre os estratos intermediários da força de trabalho, contribuindo para a explicação da tendência mais geral de polarização dos mercados de trabalho, como já apontada por vários estudiosos.

David Autor demonstra que nos EUA a participação dos empregos de baixa remuneração (como em serviços domésticos, cuidados pessoais, preparação de alimentos, vigilância) e de alta remuneração (em ocupações técnicas, profissionais e gerenciais) crescem, às expensas da participação de ocupações de nível intermediário de remuneração (como operadores de máquinas em linhas de produção e trabalhadores administrativos) (AUTOR 2014, AUTOR *et al.* 2006).

Martin Ford, tecnólogo e empresário de sucesso no Vale do Silício, argumenta que os sistemas baseados em IA são muito mais produtivos do que os seres humanos e têm potencial para substituí-los em um arco quase infinito de funções, inclusive aquelas que exigem criatividade; e ilustra sua tese com abundância de exemplos bastante eloquentes, que incluem sistemas já existentes capazes de compor sinfonias, pintar quadros e escrever romances, que podem não

conquistar prêmios de crítica, mas não se distinguem das muitas obras produzidas por seres humanos, igualmente não premiadas, que abastecem os respectivos mercados (FORD 2015).

É este tipo de diagnóstico que está na raiz das propostas de renda universal que hoje se popularizam como ideia e que já são experimentadas em países como a Suíça. Trata-se, neste caso, de substituir crescentemente o mercado de trabalho, como principal mecanismo de distribuição de renda, por um sistema de proteção social não vinculado a contribuições. A ideia é certamente atraente, mas resta saber até que ponto a fórmula suíça (um centro financeiro internacional com renda altíssima e população minúscula) pode servir de modelo para o resto do mundo.

Só o tempo dirá se a história vai ou não se repetir desta vez. O que não se discute é que a presente onda de inovações tecnológicas (diferentemente das anteriores) não afeta apenas as ocupações braçais e penosas, que - abstraído o período de adaptação, em si mesmo penoso para aqueles trabalhadores que se tornaram redundantes - não deixam saudades; ela vai no coração das ocupações e das profissões que constituem as classes médias, amplas nos países mais ricos, reduzidas nos países mais pobres, resultando em aumento das desigualdades salariais.

Todos os estudos citados até aqui tratam de países de renda alta, com mercados de trabalho largamente compostos por assalariados formais, trabalhando para empresas e instituições formais, com educação formal secundária ou superior. São também países com taxas de crescimento populacional baixa ou negativa e que historicamente recorrem à importação de mão-de-obra para o desempenho de atividades de menor qualificação na agricultura, nos serviços domésticos e de cuidados, na construção civil e nos transportes, por exemplo. Uma mão-de-obra que, muito frequentemente, não desfruta dos mesmos direitos trabalhistas e sociais que os trabalhadores nacionais, inclusive porque muitos não têm *status* legal como imigrantes. Em outras palavras, são mão-de-“obra barata e descartável” (CRAIG *et al.* 2015).

Atividades intensivas em mão-de-obra, como as indústrias de roupas e de equipamentos eletrônicos, há muito foram terceirizadas para países do terceiro mundo. As tecnologias de que estamos

falando são criadas nos países ricos e respondem a uma estrutura de oferta de força de trabalho limitada e controlada, recebendo salários relativamente altos em relação ao restante do mundo.

No contexto de ampla integração econômica entre os países em que vivemos, essas tecnologias se transferem rapidamente para os demais países do mundo, onde vão interagir com estruturas sociais e institucionais muito diferentes. Quais são os impactos, neste caso? Vamos refletir sobre essa pergunta tomando as plataformas de trabalho como exemplo.

Plataformas de trabalho

Conquanto as plataformas de trabalho representem modalidades novas e disruptivas de negócios, elas interagem com os sistemas regulatórios já existentes, ora forçando seus limites, ora se adaptando a eles, e sempre que possível explorando os espaços de sombra que todos esses sistemas contêm (VALDEZ, 2022).

As plataformas de transportes de passageiros e de entregas de mercadorias são de longe as mais numerosas. Segundo levantamento global da OIT (2020), existiam 106 plataformas de táxis e 383 de entregas. Para se ter uma ideia do ritmo de crescimento destes negócios, em 2010 a soma destas duas modalidades de plataformas mal chegava a 50 (ILO, 2021), cerca de 10 vezes menos do que em 2020. Essas são também as duas modalidades de serviços mais pesquisadas, havendo já uma grande quantidade de estudos focados, sobretudo, nas relações de trabalho que elas promovem (ROSENBLAT 2018, 2016, HUWS 2014, SRNICEK 2016, SCHOLZ 2017, SURIE 2020, PAIS 2020, MILKMAN *et al.* 2021, ABILIO 2020).

De modo geral, a literatura aponta que os trabalhadores engajados nestas atividades exercem longas jornadas, arcam com os custos e riscos inerentes ao serviço, mas estão “subordinados ao controle” exercido pelas plataformas e estão sujeitos a punições, que vão de descontos a suspensões temporárias e desligamento definitivo da plataforma. Na maioria dos países, não são legalmente enquadrados como empregados e não são cobertos pelas legislações trabalhistas que se aplicam aos assalariados formais. Como regra, as plataformas os classificam como prestadores de serviços autônomos,

livres para estabelecer seus horários de trabalho e para manter vínculos com outros contratantes, razão pela qual não lhes devem obrigações patronais. Mas a forma de enquadramento do trabalho está longe de ser pacífica e desde a criação das primeiras plataformas deste tipo, essa questão tem motivado inúmeras disputas judiciais.

Em países, como o Brasil, a escassez de empregos assalariados e a fragilidade dos instrumentos de proteção contra o desemprego e a pobreza fazem com que parcelas muito grandes da força de trabalho tenham que sobreviver regularmente por meio do chamado trabalho por conta própria, habituados a uma condição ocupacional marcada pela instabilidade e pela ausência de regras e direitos. É uma força de trabalho numerosa e disponível para relações de trabalho do tipo que as plataformas preferem. Com níveis de renda relativamente baixos, essa população, também consumidora, está altamente propensa a aderir a serviços com custos rebaixados. Do ponto de vista dos governantes, pressionados pela incapacidade de prover serviços como transportes públicos de qualidade e a preços acessíveis, o surgimento de alternativas como os aplicativos representa uma "solução" providencial, a ponto de compensar quaisquer problemas que eles possam trazer. Tudo isso inibe a implementação de regulamentações restritivas como as que são adotadas em cidades de países mais ricos, como nos exemplos de Madri e Berlim (VALDEZ 2022).

Pelo seu pioneirismo e tamanho, a Uber é sem dúvida um caso muito relevante para os estudos sobre trabalho em plataformas. O mesmo vale, de maneira mais geral, para todo o setor de transporte de passageiros e entregas de mercadorias, com características essencialmente muito semelhantes. Mas até que ponto esses setores servem de parâmetro para a enorme variedade de plataformas de trabalho que se multiplicam em outros tipos de atividade, como o uso indiscriminado da expressão "uberização do trabalho" sugere? A resposta a essa pergunta só pode ser tentativa, uma vez que as pesquisas, até aqui, se concentram muito nessas atividades específicas e ainda são escassas para outras categorias de serviços. Mas é possível trabalhar a questão por meio do contraste com outros tipos de plataformas, como as chamadas *online web-based platforms*, onde os produtos do trabalho são de natureza inteiramente digital.

A história das plataformas de trabalho tem início cerca de dez anos antes da fundação da Uber. A pioneira foi a Amazon, com seu Mechanical Turk (ILO 2021). Diferentemente das plataformas de tipo *local-based*, que capturam serviços já há muito existentes e normalmente realizados por trabalhadores autônomos (formal ou informalmente), o motor da expansão das plataformas *web-based* são as próprias inovações tecnológicas propiciadas pelas tecnologias de informação e, muito particularmente, pela difusão dos sistemas baseados em inteligência artificial. O "processo de produção" destes sistemas envolve uma extensa e complexa divisão do trabalho e o "produto final" é uma composição de muitas formas de trabalho diferentes.

As grandes plataformas desse setor, como a UpWork e a Appen, oferecem uma cesta de serviços complementares entre si, a serem executados por tipos muito diferentes de trabalhadores, desde os cientistas de dados, profissionais altamente qualificados e ainda escassos, sendo muito bem remunerados, responsáveis por transformar grandes bases de dados em produtos comerciais lucrativos, até os trabalhadores anônimos das chamadas fazendas de cliques, que em troca de centavos realizam tarefas repetitivas e maçantes, porém, indispensáveis para o treinamento dos algoritmos, passando por engenheiros e arquitetos de bancos de dados, programadores, desenvolvedores, especialistas em segurança cibernética, entre muitas outras especialidades.

A diversidade e a heterogeneidade da força de trabalho necessária exigem formas de recrutamento e seleção, contratação e remuneração também muito variadas. Nesse contexto, encontramos contratos por projetos, por resultados, por pacotes de serviços ou tarefas, conforme a complexidade dos produtos; encontramos consultores e *freelances*, recebendo por hora-trabalho; entre programadores são comuns as competições, das quais muitos participam, mas apenas o ganhador ou ganhadores recebem "prêmios" em dinheiro; encontramos testadores de jogos, que nada recebem; e, mais na ponta, encontramos também sistemas de trabalho em "multidões" (*crowd-work*) em que a força de trabalho é recrutada a granel e cada um é pago pelo volume de tarefas realizados.

Essas plataformas fornecem também os meios técnicos para a concatenação dos vários trabalhos necessários para a montagem dos produtos e os instrumentos de controle e gestão dessa força de trabalho. Elas formam ecossistemas complexos de trabalho, por sua vez, conectados com outros sistemas complexos, necessários ao seu próprio funcionamento, como as grandes plataformas de computação em nuvem e armazenamento de bancos de dados, as plataformas de recrutamento de profissionais, as empresas de segurança cibernética, as fazendas de cliques, as *startups* responsáveis por novas ferramentas e produtos e centros de pesquisa.

Por se tratar de trabalho digital, a relação destas atividades com o espaço físico (e com as autoridades que o regem) é inteiramente diferente da que caracteriza os serviços de tipo *local-based*. Nestes últimos, tem-se como características: a) os consumidores e os prestadores de serviços que as plataformas conectam, como regra, encontram-se em uma mesma circunscrição geográfica, geralmente um município; b) as empresas têm que negociar os seus termos de funcionamento com autoridades claramente estabelecidas e dotadas de poderes suficientes até mesmo para bani-las; c) cada trabalhador realiza sozinho o serviço demandado pelo consumidor.

Diferentemente disso, as plataformas de trabalho digital oferecem produtos que são partes componentes de outros produtos; algumas organizam-se como ecossistemas que podem conter cadeias completas de produção, com toda a divisão de trabalho que lhe corresponde; outras especializam-se em partes destas cadeias.

De modo muito semelhante ao que ocorre em cadeias globais de bens manufaturados, em que as grandes empresas que as comandam segmentam e dispersam geograficamente as várias etapas de produção, buscando situar cada atividade onde ela possa ser realizada com menores custos (logística, salários, tributários etc.), o trabalho digital se presta, de uma forma muito mais radical, à dispersão geográfica. Diferentemente do que se passa na manufatura, cuja dispersão depende de várias infraestruturas (ferrovias, estradas, portos, navios), enfrenta barreiras tributárias e tem custos de gerenciamento de fornecedores e subcontratados, a integração de cadeias digitais de trabalho depende de uma única infraestrutura, que

são os meios de comunicação digitais. Eles são suficientes para que empresas dotadas de sistemas complexos de informação e poder computacional superlativo (ou seja, muito conhecimento científico e muito capital), instaladas em alguns poucos edifícios e empregando diretamente uma força de trabalho muito menor do que qualquer indústria de porte, sejam capazes de reunir, instantaneamente, o trabalho dos cientistas da computação, que trabalham no Vale do Silício e são regiadamente remunerados, os programadores que vivem na Ucrânia, onde engenheiros recebem salários ínfimos na comparação com os EUA e de uma massa anônima de clicadores espalhados pela Índia ou Paquistão, cujo trabalho praticamente não tem valor, pelo menos não na escala de valores monetários em que operam as grandes corporações.

As plataformas do tipo *web-based* são as infraestruturas que sustentam essas cadeias de trabalho, as quais são usufruídas principalmente por outras empresas que fazem uso intensivo das mesmas tecnologias digitais para realizar seus negócios, quaisquer que sejam. Essa infraestrutura, no entanto, lhes serve como fonte de poder não tanto para superar limites regulatórios, porque não há autoridade que se sobreponha a estes tipos de arranjo, mas sobretudo para capturar valor. Isso já era verdade para as cadeias de manufatura, mas o trabalho digital vai muito mais longe na liberdade que confere às empresas na hora de mobilizá-lo.

Para encerrar este ponto, cumpre destacar que o que há de comum entre estes dois grandes tipos de plataformas de trabalho é que elas dependem, pelo menos em parte, da existência de grandes contingentes de força de trabalho disponíveis para vender trabalho de forma fragmentária, seja porque estão desempregados e não dispõem de seguro-desemprego ou outro sistema de proteção social, seja porque não tem capital para fazer com que seus pequenos negócios produzam uma renda suficiente, seja porque sempre sobreviveram por meio de trabalhos temporários e bicos, seja porque os salários dos empregos disponíveis são muito baixos, seja porque a obrigação de cuidar de parentes ou problemas de saúde as impedem de assumir um emprego em tempo integral. As fontes de vulnerabilidade são muitas, mas convergem para a falta de alternativas. E, felizmente para as empresas, são abundantes as populações que se enquadram nessas

situações, embora sua distribuição pelo mundo não seja homogênea, concentrando-se principalmente nos países menos desenvolvidos.

Conclusões

A onda de inovações tecnológicas que estamos vivendo impacta de forma significativa no mercado de trabalho, não necessariamente eliminando empregos em massa, mas alterando sua natureza.

Os estudos sobre os EUA sugerem que muitas das ocupações situadas no meio da pirâmide ocupacional, normalmente associadas a empregos estáveis e bem remunerados, cobertos por direitos sociais e trabalhistas e benefícios que as empresas custeiam, são afetadas negativamente. Já as profissões ligadas a essas tecnologias se veem valorizadas em função da escassez relativa de profissionais, mas isso não deve durar para sempre, porque empresas, governos e as pessoas que estão em posição de se qualificar estão trabalhando para preencher estas lacunas.

Já em países de nível intermediário de desenvolvimento, como o Brasil, os níveis agudos de escassez de profissionais especializados em novas tecnologias de informação não apenas inflacionam seus salários, mas retardam a difusão dessas tecnologias, aumentando o *gap* entre as empresas aqui instaladas e as de países que estão liderando essas inovações. Além disso, com a expansão das empresas plataformas, nada impede que a força de trabalho com os requisitos necessários e disponíveis para o trabalho digital, em seus diversos níveis de complexidade, do engenheiro desempregado que passa a realizar tarefas de programação, ao estudante sem renda que se socorre em fazendas de cliques, seja mobilizada por empresas estrangeiras.

Essas tendências apontam para um possível processo de desassalariamento, ou seja, de desvinculação entre trabalho e sistema de proteção social. Aqui, como na maior parte do chamado mundo em desenvolvimento, esse vínculo sempre foi limitado, cobrindo apenas parcialmente a força de trabalho. A crença em que o desenvolvimento econômico trataria de remediar a situação, eliminando as formas mais

“arcaicas” de relações de trabalho, há muito deixou de ser convincente.

Para um país como o Brasil, o desafio é duplo: por um lado, é indispensável desenvolver capacidades próprias nos vastos domínios dessas novas tecnologias, de modo a garantir que elas tenham aplicações que beneficiem o País como um todo e não apenas as empresas envolvidas; por outro, repensar o sistema de proteção social - particularmente os seus mecanismos de financiamento - de modo a que toda a população tenha acesso a ele, independentemente das relações de trabalho em que está inserida. De quebra, isso resolveria também problemas muito mais antigos do que os que a atual revolução tecnológica pode vir a causar.

Referências Bibliográficas

ABÍLIO, L.C. (2020) Uberização: a era do trabalhador just-in-time? *Estudos Avançados* 34, nº 98.

AUTOR, D. (2014) **Polanyi's Paradox and the Shape of Employment Growth**. Cambridge-MA: National Bureau of Economic Research. <https://doi.org/10.3386/w20485>.

AUTOR, D.H.; KATZ, L.F.; KEARNEY, M.S. (2006) The Polarization of the U.S. Labor Market. *AEA Papers and Proceedings* 96(2).

AUTOR, D. (2020) **The Work of the Future: Building Better Jobs in an Age of Intelligent Machines**. MIT: MIT Work of the Future Task Force.

BRYNJOLFSSON, E.; McAFFEE, A. (2011) **Race Against the Machine**: How the Digital Revolution Is Accelerating Innovation, Driving Productivity, and Irreversibly Transforming Employment and the Economy. Lexington: Digital Frontier Press.

BRYNJOLFSSON, E.; MITCHELL, T.; ROCK, D. (2018) What Can Machines Learn and What Does It Mean for Occupations and the Economy? *AEA Papers and Proceedings* 108: 43-47. <https://doi.org/10.1257/pandp.20181019>.

CRAIG, Gary. WAITE, Louise. LEWIS, Hannah. SKRIVANKOVA, Klara.(2015) Vulnerability. *Exploitation and Migrants: Insecure Work in a Globalised Economy*. New York: Palgrave Macmillan.

DUNAWAY, W.A. (2013) **Gendered Commodity Chains: Seeing Women's Work and Households in Global Production**. Stanford: Stanford University Press.

FORD, M. (2015) **Rise of the Robots**: Technology and the Threat of a Jobless Future. New York: Basic Books.

-
- HUWS, U. (2014) **Labor in the Global Digital Economy**: The Cybertariat Comes of Age. New York: NYU Press.
- ILO - International Labour Office (2021) **World Employment and Social Outlook 2021**: The role of digital labour platforms in transforming the world of work. Geneva.
- LANE, M.; SAINT-MARTIN, A. (2021) The Impact of Artificial Intelligence on the Labour Market: What Do We Know so Far? *OECD Social, Employment and Migration Working Papers* 256. <https://doi.org/10.1787/7c895724-en>.
- MILKMAN, R.; ELLIOTT-NEGRI, L; GRIESBACH, K.; REICH, A. (2021) Gender, Class, and the Gig Economy: The Case of Platform-Based Food Delivery. *Critical Sociology* 47(3) <https://doi.org/10.1177/0896920520949631>.
- PAIS, I. (2020) Digital Platforms and the Transformations in the Division of Labor. In: *Handbook of Labor, Human Resources and Population Economics*. Springer International Publishing. https://doi.org/10.1007/978-3-319-57365-6_16-1.
- ROSENBLAT, A. (2016) **Algorithmic Labor and Information Asymmetries**: A Case Study of Uber's Drivers. California: Internacional Jornal Communication. <https://ijoc.org/index.php/ijoc/article/view/4892/1739>
- ROSENBLAT, A. (2018) **Uberland**: How Algorithms Are Rewriting the Rules of Work. California: University of California Press.
- SCHOLZ, T. (2017) **Überworked and Underpaid**: How Workers Are Disrupting the Digital Economy. Cambridge: Polity Press.
- SURIE, A. (2020) On-Demand Platforms and Pricing: How Platforms Can Impact the Informal Urban Economy, Evidence from Bengaluru, India. *Work Organisation, Labour & Globalisation* 14(1). <https://doi.org/10.13169/workorglaboglob.14.1.0083>.
- SRNICEK, N. (2016) **Platform Capitalism**. Cambridge: Polity Press.
- VALDEZ, J. (2022) The Politics of Uber: Infrastructural Power in the United States and Europe. In: *Regulation & Governance*. <https://doi.org/10.1111/rego.12456>
-

Autor

Alvaro A Comin

Professor do Departamento de Sociologia da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo desde 1999 e pesquisador do Centro de Inteligência Artificial da USP (C4AI-USP), onde coordena o grupo de pesquisa “Mercados Digitais de Trabalho”.



Créditos da imagens-base das ilustrações:

Fotografias em domínio público. Fonte: Wikimedia Commons (<https://>

Capa:

"U.S. Army Photo", from M. Weik, "The ENIAC Story" A technician changes a tube. Caption reads "Replacing a bad tube meant checking among ENIAC's 19,000 possibilities."

https://commons.wikimedia.org/wiki/File:ENIAC-changing_a_tube.jpg

Verso da capa:

Two of the ENIAC programmers, are preparing the computer for Demonstration Day in February 1946. "U.S. Army Photo" from the archives of the ARL Technical Library

[https://commons.wikimedia.org/wiki/File:Two_women_operating_ENIAC_\(full_resolution\).jpg](https://commons.wikimedia.org/wiki/File:Two_women_operating_ENIAC_(full_resolution).jpg)

P. 5

The ENIAC was the first electronic digital computer.

https://commons.wikimedia.org/wiki/File:Eniac_Aberdeen.jpg

P. 12

Photograph of World's First Computer, the Electronic Numerical Integrator and Calculator

[https://commons.wikimedia.org/wiki/File:World's_First_Computer,_the_Electronic_Numerical_Integrator_and_Calculator_\(ENIAC\).gif](https://commons.wikimedia.org/wiki/File:World's_First_Computer,_the_Electronic_Numerical_Integrator_and_Calculator_(ENIAC).gif)

P. 179

ENIAC (Electronic Numerical Integrator And Computer) in Philadelphia, Pennsylvania

https://commons.wikimedia.org/wiki/File:Glen_Beck_and_Betty_Snyder_program_the_ENIAC_in_building_328_at_the_Ballistic_Research_Laboratory.jpg