UNIVERSIDADE DE LISBOA

FACULDADE DE DIREITO

A PROTEÇÃO DE DADOS PESSOAIS E A REGULAÇÃO DA INTELIGÊNCIA ARTIFICIAL: EM DIREÇÃO A UM CONSTITUCIONALISMO DIGITAL.

FELIPE MÜLLER DORNELAS

LISBOA

2024

FELIPE MÜLLER DORNELAS

Aluno n.º 63678

A PROTEÇÃO DE DADOS PESSOAIS E A REGULAÇÃO DA INTELIGÊNCIA ARTIFICIAL: EM DIREÇÃO A UM CONSTITUCIONALISMO DIGITAL.

Dissertação apresentada como exigência parcial à obtenção do título de Mestre em Direito e Ciência Jurídica ao Curso de Mestrado em Direito e Ciência Jurídica, na Especialidade de Direito Constitucional da Faculdade de Direito da Universidade de Lisboa, sob a orientação do Senhor Professor Doutor Pedro Alexandre Vicente de Araújo Lomba.

LISBOA

2024

DEDICATÓRIA

Dedico esta dissertação, em primeiro lugar à Deus, pois a cada dia tenho mais certeza de suas bençãos em minha vida e de minha família! Olhando em retrospectiva todos os acontecimentos que me levaram a frequentar o tão sonhado curso de mestrado na melhor universidade de Portugal, com excelentes professores e estrutura única, numa cidade que passou a ser meu local preferido no mundo, só foi possível porque Ele tornou realidade. Obrigado, Senhor!

De tamanha importância na minha vida é minha esposa, para quem agradeço em segundo lugar, por ter, desde o início, acreditado nesse projeto e embarcado no sonho, sem momento algum duvidar que era possível e, ainda, por me dar todo o suporte emocional e afetivo que precisei para chegar até este momento. Obrigado por todo amor, companheirismo e amizade, sem você não teria conseguido! Te amo, Rê!

Aos meus pais eu agradeço imensamente a criação que me deram, fundamentalmente os princípios que me passaram, que funcionam como minha fundação e meu guia, habilitando-me a ter estrutura para enfrentar todos os desafios da vida, especialmente este de cursar o mestrado fora do país. Obrigado por tudo, mãe e pai! Amo vocês!

Às minhas irmãs, Nathália e Victória, por serem as melhores irmãs, companheiras e acreditarem em mim, mais do que eu, em muitas vezes! *Liebe euch alle*! *vielen Dank*!

Agradeço aos familiares, colegas da FDUL e aos amigos que, direta ou indiretamente, contribuiriam para este momento, em especial ao Dom Eurico dos Santos Veloso (*in memoriam*) que, sem sombra de dúvidas, abriu a porta para realização deste sonho!

E, ao meu filho, Benjamin: obrigado simplesmente por existir! Papai te ama!

AGRADECIMENTOS

Meu agradecimento especial aos professores que compartilharam tão humildemente seus conhecimentos comigo durante todo o mestrado e possibilitaram meu crescimento acadêmico em muitas maneiras, meu muito obrigado!

Todavia, preciso destacar aqui alguns ensinamentos que fogem ao Direito, pois são verdadeiras lições de vida:

Em Justiça Constitucional, desde o primeiro instante soube que o Prof. Doutor Pedro Lomba seria meu orientador, pois, para além de todo o conhecimento jurídico impecável, percebia que sua visão de vida era muito parecida com a minha, o que gerou uma empatia e a certeza que seria o melhor professor para me conduzir até este momento. Obrigado!

Em Direitos Fundamentais, tive uma grande lição de humildade e profissionalismo ao ver o Prof. Doutor Catedrático Jorge Miranda dedicar-se às aulas com grande entusiasmo e paixão, ainda que durante todo o período de COVID-19, sem temer estar presente com os alunos para exercer seu magistério.

Em Direito Constitucional, deparei-me com o que é precisão, profundidade e inteligência jurídica ímpar, ao ser aluno do Prof. Doutor Reis Novais, mas sobretudo com seu vigor e profissionalismo na docência.

Em Metodologia Científica, aprendi com o Prof. Doutor Pedro Sanches que idade cronológica não é requisito para ser um grande pesquisador ou profissional; que vale mais o esforço aliado a técnica.

E, por fim, gostava de agradecer, igualmente, aos colegas do Lisbon Public Law (LPL), na pessoa do Prof. Doutor Domingos Farinho, a quem me recepcionou como pesquisador no projeto Lisbon Digital Rights and Freedoms, onde pude aprofundar sobremaneira meus conhecimentos, possibilitando de várias formas o resultado final também desta dissertação.

EPÍGRAFE

Domine, non sum dignus, ut intres sub tectum meum: sed tantum dic verbo, et sanabitur anima mea (Mt. 8:8).

RESUMO

O avanço da tecnologia interfere de maneira significativa em todos os aspectos da vida cotidiana da sociedade, desde assuntos mais comezinhos até questões políticas de grande relevância, em alguns casos desaguando em revoluções. Passamos da antiga esfera pública tradicional para o surgimento da nova praça pública digital das plataformas digitais, especialmente diante do crescimento vertiginoso de tecnologias como os sistemas de inteligência artificial e o uso massivo de dados pessoais para alavancar e propiciar a continuidade e avanço desta. Com isso, novos desafios surgiram e o direito passou a não mais dar respostas adequadas com os antigos instrumentos jurídicos, necessitando de uma nova abordagem que conseguisse compreender o fenômeno social e apresentar respostas coerentes, momento em que surge o constitucionalismo digital, objetivando endereçar tais respostas, especialmente focado na proteção de direitos fundamentais no ambiente digital e na regulação do mesmo a partir de preceitos fundantes do Estado de Direito.

PALAVRAS-CHAVE

Constitucionalismo Digital; Proteção de Dados Pessoais; Inteligência Artificial; Regulação; Plataformas Digitais

ABSTRACT

The technological progress and breakthrough significantly interferes in all aspects of society's daily life, from the most trivial matters to political issues of great relevance, in some cases leading to revolutions. We have moved from the old traditional public sphere to the emergence of the new digital public square of digital platforms, especially given the dizzying growth of technologies such as artificial intelligence systems and the massive use of personal data to leverage and promote its continuity and development. As a result, new challenges emerged and the law no longer provided adequate responses with the old legal instruments, requiring a new approach that could understand the social phenomenon and present coherent responses, at which point the digital constitutionalism thesis emerged, aiming to address such responses, especially focused on the protection of fundamental rights in the cyberspace and its regulation based on the founding precepts of the Rule of Law.

KEYWORDS

Digital Constitutionalism; Personal Data Protection; Artificial Intelligence; Regulation; Digital Platforms

NOTA DE ADVERTÊNCIA

A presente dissertação de mestrado foi escrita na Língua Portuguesa segundo seu uso no Brasil, conforme as normas do Novo Acordo Ortográfico da Língua Portuguesa.

Para a formatação, estilo e citação foram seguidas as diretrizes da Associação Brasileira de Normas Técnicas – ABNT 10520:2023.

LISTA DE SIGLAS E ABREVIATURAS

ADI	Ação Direta de Inconstitucionalidade
AIA	*Artifcial Intelligence Act*
ANPD	Agência Nacional de Proteção de Dados brasileira
BDSG	*Bundesdatenschutzgesetz*
BvG	*Bundesverfassungsgericht*
CNNs	Redes Neurais Convolucionais
CRFB/1988	Constituição da República Federativa do Brasil
DL	*Deep Learning*
EC	Emenda Constitucional
ed., eds.	Edição, edições; editora, editoras
et al.	*et alii*
EU	União Europeia
FDUL	Faculdade de Direito da Universidade de Lisboa
GANs	Redes Generativas Adversariais
GPT	Generative Pre-trained Transformer
IA	Inteligência Artificial
LGPD	Lei Geral de Proteção de Dados
LINDB	Lei de Introdução às Normas do Direito Brasileiro
LLMs	*Large Language Models*
ML	*Machine Learning*
n.	Número
PLIA	Projeto de Lei n.º 2338/23
PLN	*Natural Language Processing*
rev.	Revisão, revista
RGPD	Regulamento Geral de Proteção de Dados
RIPD	Relatório de Impacto à Proteção de Dados Pessoais
RNNs	Redes Neurais Recorrentes
RSD	Regulamento dos Serviços Digitais
SF	Senado Federal
STF	Supremo Tribunal Federal
TFUE	Tratado sobre o Funcionamento da União Europeia
v.	Volume
VAEs	*Variate Autoencoders*

SUMÁRIO

A PROTEÇÃO DE DADOS PESSOAIS E A REGULAÇÃO DA INTELIGÊNCIA ARTIFICIAL: EM DIREÇÃO A UM CONSTITUCIONALISMO DIGITAL.

INTRODUÇÃO

Os influxos da sociedade da informação, notadamente marcada pela digitalização das relações sociais, produziu no Direito Constitucional uma corrente de pensamento denominada de *constitucionalismo digital,* que objetiva compreender o espaço digital e desenhar respostas adequadas que englobem os principais gargalos do mundo digital.

Para Mendes e Fernandes (2020), estes desafios albergam: i) reconhecimento de novos direitos fundamentais; ii) limitação da capacidade de violação de direitos fundamentais na internet; iii) novas formas de controle social sobre os atores do espaço, tanto governamentais, quanto privados; iv) a velocidade de resposta aos problemas; v) a não limitação aos territórios soberanos e o acúmulo de poder por agentes privados (Mendes e Fernandes, 2020.

Para EDOARDO CELESTE (2019) o constitucionalismo digital trata-se de uma verdadeira ideologia constitucional em que se estrutura em um quadro normativo de proteção dos direitos fundamentais e de reequilíbrio de poderes na governança do ambiente digital.

Assim, ao mesmo tempo que a internet amplificou o acesso à informação e promoveu uma abertura aos direitos fundamentais e ao exercício da cidadania e da democracia, também demonstrou sua faceta autoritária, vigilante e manipuladora de comportamentos dos usuários, tanto de governos, quanto pelo setor privado. À evidência, um espaço que foi concebido para ser um *locus* de liberdade quase que absoluta, hoje é repensado para se adequar aos princípios constitucionais estruturantes do Estado de Direito e a regulação é uma ideia que se mostra no centro do debate atual.

Portanto, diante dessas transformações impostas ao regime de proteção de direitos fundamentais que, nos últimos anos, estados nacionais,

entidades privadas e organizações sociais têm se mobilizado para reestabelecer o equilíbrio constitucional nos espaços digitais ().

Consequentemente, dentro deste escopo teórico indicado acima, pretendemos investigar, a partir de um enquadramento teórico-jurídico do *Constitucionalismo Digital,* quais os endereçamentos aos desafios existentes e ainda o porvir do ciberespaço.

Para tanto, iniciamos o primeiro capítulo com os conceitos sobre o que seria o *Constitucionalismo Digital*, obras e autores mais renomados sobre o assunto, sua origem, sua aplicação e a transmutação da esfera pública tradicional às plataformas digitais, chamada de “esfera pública digital”, bem como uma introdução ao papel do dados pessoais e sua imbricada conexão com os sistemas de inteligência artificial, a fim de desnudar o panorama geral da investigação e munir o leitor com substratos iniciais para compreender o desenrolar dos itens seguinte.

Em seguida, no capítulo 2, iniciamos o tratamento do assunto acerca da inteligência artificial, trazendo conceitos e o cenário histórico, presente e futuro desta tecnologia na vida humana, especialmente seus impactos de ordem jurídico-legal. Para tanto, deve-se perpassar igualmente pelos benefícios, riscos e modelos de sistemas de IA existentes hodiernamente.

Uma vez exposto o grande quadro teórico do Constitucionalismo Digital e, adentrado à inteligência artificial, faz-se inevitável correlacionar a questão de dados pessoais e sua proteção com o dinâmica dos sistemas de IA partindo de um breve histórico e evolução do direito à proteção de dados pessoais; seu regime jurídico no Brasil e União Europeia e interseções ínsitas aos dois temas.

Por consequência, a pesquisa indica aprofundar o estudo da proteção de dados pessoais e Inteligência Artificial (IA) para englobar, compreender e responder os questionamentos atuais sobre um determinado tipo de IA, qual seja, a de natureza generativa. Logo, o terreno ainda pouco explorado sobre IA generativa e seus potenciais benéficos e maléficos à sociedade obriga a abordagem deste subtema, onde explanaremos uma breve introdução aos conceitos, o funcionamento e aos modelos existentes, para tentar perceber de

que forma a proteção de dados pessoais poderá auxiliar na tutela deste novo cenário tecnológico.

Por fim, trataremos da regulação entre Brasil e União Europeia, nomeadamente através dos principais diplomas sobre o tema proteção de dados pessoais e inteligência artificial, a fim de compreender as respostas possíveis do Constitucionalismo Digital.

1) Constitucionalismo Digital:

1.1) Origem, conceito e enquadramento teórico-jurídico:

Vivenciamos e experimentamos o fenômeno da digitalização de todos os aspectos da vida, em magnitude total no que toca a persuasão deste contexto nas relações sociais, interpessoais, econômicas, jurídicas e políticas em que, ao mesmo tempo proporcionam grandes benefícios à humanidade, mas igualmente implicam desafios e problemas até então não experimentados, de difícil enfrentamento e endereçamento.

A evolução tecnológica proporcionou uma escalabilidade no acesso à informação que, por sua vez, transformou o mundo que conhecíamos num novo espaço, onde os ambientes on-line, leia-se plataformas digitais, revelaram-se o *locus* da antiga praça pública, onde se faz o debate público sobre os assuntos mais importantes e se consolida a opinião pública e os interesses maiores da sociedade.

A renomada socióloga de Harvard, SHOSHANA ZUBOFF (2020) é de precisão peculiar ao perceber que a informação é o bem mais precioso nesse novo sistema envolto ao ciberespaço, cunhando tal momento de “Capitalismo de Vigilância”, pela estreita relação com as fases de evolução do capitalismo, fulcrado na melhoria da eficiência e produtividade dos processos de troca de informação ou dados, mas também gerador de um sistema de vigilância sem precedentes.

Corroborando este contexto, MANUELL CASTELLS (1999. P. 458) diz que “esta nova economia surge no final do século XX, levada pelas mudanças

tecnológicas e de transmissão de informação, que conferiu a base material indispensável para o seu nascimento". Assim, o alcance global e instantâneo da informação ao redor do mundo proporcionou a disrupção para um novo estilo de vida conectado ao espaço cibernético.

Destarte, o jurista RICARDO CAMPOS (2022) é categórico sobre a abrangência desta nova era, para quem:

> "[...] o mundo digital cria não apenas uma nova ordem de conhecimento, mas também uma nova ordem de cultura, ao decantar e depurar a antiga forma de conhecimento, que era orientada para organizações, substituindo a velha forma de economia, baseada na revolução industrial, prevalecente da metade do século XIX até ao século XX, por uma nova forma de economia, consubstanciada em dados e, lastreada em plataformas. (CAMPOS, 2022, p. 277)

Em adição, o suprarreferido jurista afirma que tal transição de ordens, dada sua novidade e ausência de precedentes, pode oferecer abordagens extremas, tanto autoritárias, quanto demasiadamente liberais, ou seja, endereçamentos desproporcionais aos desafios que se apresentam:

> "Este processo de transição para uma nova ordem digital de conhecimento e cultura, no qual a 'não-conceitualidade' (*Unbegriefflichkeit*) existente do novo (*Blumenberg*) desafia a antiga ordem baseada na organização, ao mesmo tempo em que reduz, por um lado, simples noções de restauração de uma "comunidade" que pode sempre assumir características autoritárias, não liberais e, por outro lado, soluções libertárias que preferem uma abordagem na qual '*anything goes*'". (CAMPOS, 2022, p. 278)

Neste ambiente do ciberespaço nasce um maior protagonismo de atores privados, detentores das tecnologias que norteiam a vida no espaço digital, nomeadamente via plataformas digitais, que passaram a ter enorme poder econômico, social e político, exercendo, inclusive, a normatização e interferência em direitos fundamentais de utilizadores, especialmente aqueles ligados à liberdade de expressão, autonomia privada, privacidade e proteção de dados pessoais.

Este oligopólio, formado pelas maiores empresas de tecnologia, opera sob um modelo de negócios baseado em utilização de algoritmos e dados

pessoais, especialmente desenhado para rentabilizar a partir dados oriundos de *profiling* de utilizadores e estimular uma permanência maior dentro da plataforma, criando-se um espaço de capitalismo baseado em vigilância.

Portanto, passamos a perceber nova concepção de sociedade, fulcrada em torno de plataformas digitais -e no mundo onde a informação é tida como um pilar de desenvolvimento e poder-, entendida também a partir de um paralelismo com a evolução das fases do capitalismo, denominada de "sociedade algorítmica" (Schuilenburg; Peeters, 2020)

Contudo, para dar um panorama resumido de como se chegou a este estágio atual de coisas, precisa-se retornar, ao menos, ao início da era capitalista. Brevemente, a história do sistema capitalista é faseada com base em três etapas: *i*) comercial; *ii)* industrial; e *iii)* financeira. Por sua vez, alguns autores afirmam existir uma quarta fase: o "capitalismo informacional"— termo desenvolvido por MANUELL CASTELLS (1999), em sua obra "A sociedade em rede".

O capitalismo de cunho comercial alavancou-se no início da formação do sistema capitalista e a consequente expansão do comércio internacional no contexto da Europa. Essa fase ficou marcada pela expansão marítima comercial e colonial, com a formação de colônias europeias em várias partes do mundo, com destaque para as Américas e para o continente africano. Nesse período, intensificou-se a prática do mercantilismo, um sistema econômico geralmente concebido como "um conjunto de práticas" não planejadas.

A segunda fase do capitalismo é chamada de "capitalismo industrial", por ter sido um efeito direto da emergência, expansão e centralidade exercida pelas fábricas graças ao processo da Revolução Industrial, iniciado em meados do século XVIII, especialmente no Reino Unido.

Por conseguinte, o dito "capitalismo financeiro" é caracterizado pelo protagonismo exercido pela especulação financeira e pela primazia em torno da bolsa de valores, que passou a ser uma espécie de "termômetro" da economia de um país. Basicamente, essa fase do capitalismo estrutura-se com a formação do mercado de ações e a sua especulação em termos de valores, taxas, juros e outros.

A atual fase, ao seu turno, é conhecida, dentre outras alcunhas, como “capitalismo informacional”, “capitalismo da vigilância” ou “quarta revolução industrial” (Schwab, 2016), designando estas expressões as tecnologias para automação e troca de dados e informações utilizando-se de conceitos como “sistemas ciber-físicos”, “internet das coisas”, “inteligência artificial”, “computação em nuvem”, dentre outros que integram o novo vocabulário do mundo digital. Portanto, o foco da “Quarta Revolução Industrial” é a melhoria da eficiência e produtividade dos processos de troca de informação ou dados, que impulsiona um capitalismo forte na recolha, processamento, tratamento e comercialização de dados (Zuboff, 2020).

Deveras, a informação sempre desempenhou papel fundamental na sociedade. O direito, como ciência social aplicada, tem a função de reconhecer e perceber esses fenômenos, para regulá-los naquilo que for necessário, em nome da justiça, segurança jurídica e da paz social. Logo, o tema ganhou novos contornos por meio do enorme avanço da tecnologia, que permite a circulação de dados pessoais e informação em tempo instantâneo.

Na economia digital, dados e informações são ativos fundamentais que podem ser considerados matérias-primas das quais o seu processamento pode gerar valor. Mesmo dados simples, quando processados com propósito específico e misturado com outras informações, pode fornecer modelos e respostas preditivas. Essas oportunidades levaram ao surgimento de novas aplicações e modelos de negócios numa nova fase do capitalismo (Policcino; De Gregorio, 2021).

Assistimos, por exemplo, a disseminação e democratização da internet, o crescimento das redes sociais, o surgimento dos *smartphones e,* principalmente, o surgimento de novos atores, em sua maioria particulares com enormes poderes, como as grandes empresas de tecnologia ou *big tech*. Contudo, os problemas advindos desse “mundo digital”, como acesso às tecnologias, desinformação, *fake news*, aumento da polarização política, discurso de ódio, discriminação nas redes sociais, malbaratamento de dados pessoais e falta de privacidade tomaram o centro das preocupações dos usuários e da sociedade.

É consabido que, cada dia mais, pessoas estão conectadas umas às outras, trocando informações de todos os tipos, o que leva a uma abertura para vigilância dos governos e de particulares donos das maiores plataformas digitais de troca de informações, onde essas interações digitais têm lugar, cada qual objetivando seus interesses, sejam eles legítimos e alinhados ao Estado Democrático de Direito, ou ilegítimos e contrários aos direitos fundamentais.

Desse modo, desde o final do século XX, a vida quotidiana tornou-se cada vez mais digital em direção a uma dimensão “on-life” (De Gregorio, 2022). Ou seja, os indivíduos exercem cada vez mais seus direitos e liberdades em um ambiente digital onipresente, onde as relações sociais são mediadas por uma mistura de entidades que expressam formas da autoridade pública e da ordem privada.

As tecnologias sempre levaram a pontos de inflexão na sociedade. No passado, as tecnologias desenvolvimentos abriram a porta para novas fases de crescimento e mudança, ao mesmo tempo que influenciou valores e princípios sociais. Logo, as tecnologias algorítmicas se ajustam dentro deste quadro. Estas tecnologias contribuíram para a introdução de novas maneiras de processar grandes quantidades de dados.

Embora estas tecnologias tenham efeitos positivos em toda a sociedade, como o de aumentar a capacidade dos indivíduos para exercerem direitos e liberdades, também levou a novos desafios constitucionais. As oportunidades oferecidas pela sociedade algorítmica necessariamente levam a uma colisão entre a tecnologia, especialmente sua opacidade e falta de responsabilização, no que foi definido como uma “algocracia” por JOHN DANAHER (2016).

Assim, não é coincidência que a transparência esteja no topo núcleo do debate sobre algoritmos, pois existem riscos para os direitos fundamentais e democracia inerente à falta de transparência sobre o funcionamento de sistemas automatizados e processos de tomada de decisão As implicações decorrentes do uso de algoritmos pode ter consequências nos direitos fundamentais dos indivíduos, como o direito à autodeterminação pessoal e informacional, liberdade de expressão e privacidade (Policcino; De Gregorio, 2019).

Em vista deste quadro geral do estado de coisas, uma corrente própria de pensamento surgiu, denominada de “Constitucionalismo Digital”, que objetiva dar direcionamento de envergadura constitucional aos desafios postos pela sociedade algorítmica, nomeadamente o *enforcement* de direitos fundamentais nas relações horizontais, o respeito à dignidade humana e ao Estado de Direito.

Outrossim, imperioso destacar que este pensamento constitucional de reação aos abusos de poderes das plataformas ocorreu incialmente na Europa., dada a sensibilidade peculiar do constitucionalismo europeu que não tolera abusos de direitos e visa proteger os direitos humanos, notadamente a dignidade da pessoa humana.

Desde os horrores da Segunda Guerra Mundial, os estados europeus começaram a incorporar e codificar a dignidade humana como princípio estruturante do Estado de Direito, elevando-o a pedra angular do Estado constitucional do pós-guerra”.

A dignidade da pessoa humana não é um conceito isolado, mas um princípio fundamental ligado aos valores e aspirações que moldam o constitucionalismo europeu. Também impulsionada pelo quadro internacional, a dignidade humana começou a emancipar o lado oriental do Atlântico do lado ocidental, onde a marca liberal do direito constitucional ainda continua a ser o fundamento principal dos direitos e liberdades fundamentais (De Gregorio, 2022).

Nesse sentido, como bem assevera o professor e doutrinador JORGE REIS NOVAIS,- posicionamento ao qual nos filiamos-, o princípio da dignidade da pessoa humana, por ser o mais central e edificante da ordem jurídica, deverá prevalecer sobre quaisquer outros, não admitindo cedência em caso de colisão com outros princípios:

> A dignidade da pessoa humana é, por definição constitucional, a base sobre que assenta a República e, com esse alcance, pode ser considerada o princípio supremo da ordem jurídica. Nessa qualidade, deve prevalecer sobre quaisquer outras razões, valores, bens interesses ou direitos que apontem em sentido divergente ou contrário [...] (Novais, 2018, p. 9).

E, adicionalmente, argumenta REIS NOVAIS (2018) que os princípios estruturantes dos Estados de Direito, nesta dicotomia entre norma- princípio e norma-regra*,* não atingem a potencialidade de verdadeiras estruturas se admitirem ponderação ante uma colisão, sendo que os princípios estruturantes devem valer sempre com uma natureza absoluta, merecendo ser observados e respeitados, sob pena de inconstitucionalidade*:*

> [...] os princípios estruturantes de Estado de Direito apresentam uma natureza que não é compatível com a sua consagração através de normas com a natureza de princípio, isto é, a sua prevalência ou cedência não pode ser remetida para ponderações de caso concreto; ao invés, a sua força vinculante prevalece, sempre. [...] Diferentemente, quaisquer que sejam as circunstâncias, não fazemos este tipo de juízo relativamente à aplicabilidade dos princípios estruturantes, não os ponderamos com outros bens, princípios, direitos ou valores para decidir qual deverá prevalecer; não ponderamos igualdade e proporcionalidade, ou dignidade e qualquer outro bem para apurar qual deve ceder (Novais, 2018, p.15).

Assente, pois, no princípio da dignidade humana, ordenamentos jurídicos de inspiração europeia tem a obrigação de atuar em prol de uma regulamentação das plataformas digitais – e de todo ciberespaço- a fim de conformar o abuso e excesso de poder à *rule of law*, vez que destes abusos e excessos decorrem lesões intoleráveis aos princípios estruturantes do próprio Estado de Direito.

Consoante ensina DE GREGORIO (2022), o constitucionalismo digital serve como uma tentativa de reformular o papel das democracias constitucionais na sociedade algorítmica, à medida que o atual contexto social, caracterizado por grandes plataformas digitais, situa-se entre os Estados-nação e os indivíduos comuns e, que se utilizam de algoritmos e sistemas de inteligência artificial para governar, sendo que as plataformas digitais globais, como o Facebook, a Amazon ou o TikTok, desempenham cada vez mais um papel crítico na intersecção entre a autoridade pública e as ordens privadas.

Sobre o assunto, o pesquisador italiano EDOARDO CELESTE (2019) diz-nos que o constitucionalismo digital pode ser entendido como o pensamento que pretende estabelecer e garantir a existência de um quadro normativo de proteção

de direitos fundamentais em harmonia com os poderes oriundos do ciberespaço"[1].

Para JOÃO PAULO LORDELO (2022), o fenômeno do constitucionalismo digital compreende:

> Um conjunto de iniciativas jurídicas que objetivam articular o exercício de direitos políticos, normas de governança e limitações do poder no ambiente digital, especialmente para limitar o exercício do poder por agentes privados na internet, nomeadamente numa sociedade algorítmica, em oposição à limitação estatal (Lordelo, 2022, p. 154).

Destarte, GIOVANNI DE GREGORIO (2021 p. 41) ensina-nos que o constitucionalismo digital, em suma, consiste na "disciplina dos limites do exercício de poder em uma sociedade em rede", e que esta denominação é o resultado da junção de dois termos distantes, em que o digital se refere à tecnologia em que as plataformas se baseiam, a forma como tratam dados e moderam o espaço público digital, e o constitucionalismo, que sempre conhecemos, e assenta na ideia de limitação do poder de quem o tem, evitando que o seu exercício seja realizado de forma discricionária e arbitrária.

Outrossim, em trabalho pioneiro sobre o tema do constitucionalismo digital os autores LEX GILL, DENIS REDEKER E URS GASSER identificaram quatro dimensões que permitem caracterizar um quadro normativo de constitucionalismo digital, a saber: a) o conteúdo substantivo das normas aborda questões políticas amplas e fundamentais que têm uma natureza inerentemente de caráter constitucional, nomeadamente exploram direitos -sejam coletivos ou individuais-, articulam limites de poder estatal e promovem uma série de normas de governação; b) as iniciativas dizem respeito a uma determinada comunidade política definida, seja explícita ou implicitamente; c) os princípios que estas normas promovem aspiram a um reconhecimento político formalizado e a uma legitimidade dentro dessa comunidade política; e d) os esforços em direção ao constitucionalismo digital apresentam um certo grau de abrangência (Gill; Redeker; Gasser, 2015).

[1] *"Ideology that aims to establish and guarantee the existence of a normative framework for the protection of fundamental rights and balancing of powers in the digital environment".*

Em resumo, o objetivo da corrente de pensamento advogada pelo *constitucionalismo digital* seria o de tutelar os direitos fundamentais -e a dignidade humana- dos indivíduos/usuários da "esfera pública online" e delimitar a atuação dos *atores digitais* deste ciberespaço, plasmado, nomeadamente, pelas *big tech* detentoras das maiores plataformas digitais do mercado.

Nesta nova dinâmica *on-life*, existe diversas vantagens do ponto de vista da organização institucional-estatal *v.g* democratizar e ampliar o acesso: i) à liberdade de expressão — e a real possibilidade de influenciar a tomada de decisões políticas —; ii) à informação, viabilizando um aumento do conhecimento e melhoria da vida democrática. Porém, traz igualmente algumas preocupações, especialmente: i) próprios contornos do que se pode dizer ou não nas plataformas- moderação de conteúdo; ii) uso de sistemas de inteligência artificial para determinar direitos e deveres; iii) excessos e abusos no uso de dados pessoais, principalmente para aumento arbitrário do lucro e discriminação, dentre outros.

Por isso, dado o papel assumido na esfera pública digital, as plataformas digitais passam a ser encaradas como sujeitos de direito público e, como tal, estão, sim, sob a tutela do direito constitucional e, consequentemente, das formas de limitação de poder e proteção de direitos e garantias fundamentais, corroborando a tese de um constitucionalismo afeto ao ciberespaço (Sousa, 2022).

Dessa forma, ancora-se agora não mais naquele constitucionalismo aprendido nas faculdades, dada sua incapacidade de tutelar os problemas atuais — mormente porque esses problemas se encontram na esfera pública digital —, mas no constitucionalismo adequado à nova realidade de poder das *big tech*, do uso indiscriminado de dados pessoais, da submissão à sistemas algorítmicos e a vigilância total sobre a vida dos utilizadores, escorado numa relação horizontal de eficácia dos direitos fundamentais e na limitação de poder exercido por estes novos *players*, chamado de constitucionalismo digital.

Dessa maneira, busca-se entender e perquirir quais os direitos e garantias fundamentais que são exercidos na esfera pública digital, quais devem

ser protegidos e quais comportamentos devem ser rechaçados num contexto supranacional e globalizado de atuação de grandes empresas de tecnologia.

É, pois, neste contexto social de enorme utilização de dados pessoais para comercialização, categorização, limitação de liberdade, influência e controle da privacidade, atrelado a um poder jamais visto nas mãos de burocratas estatais e plataformas digitais — ambos utilizadores de inteligência artificial — que surge uma corrente de pensamento que pretende equacionar, via direito constitucional, especialmente por meio de uma nova interpretação dos direitos fundamentais, bem como o reconhecimento de novos direitos, os problemas da digitalização das relações e da sociedade, denominado de constitucionalismo digital.

Logo, os poderes públicos já não são os únicos capazes de propor regulamentações, vez que são apenas uma parte da estrutura fragmentada da governança on-line, especialmente pelo surgimento de novos atores expressando seus poderes. Assim, a acumulação de poder aumenta a assimetria entre os particulares envolvidos, quais seja, big tech de um lado e usuários de outro, desaguando muitas vezes no exercício arbitrário do poder das plataformas.

1.2) Da esfera pública tradicional às plataformas digitais: a era da “esfera pública digital”:

Consoante afirma o jurista português SIMÃO SOUSA (2022), houve uma transferência da esfera pública tradicional para as plataformas digitais, ou seja, há um novo *locus* público de debate na sociedade digital, onde são exercidos diversos direitos fundamentais:

> É na esfera pública da vida em sociedade que o indivíduo institucionaliza a relação com os demais, especialmente através da liberdade de expressão. Na nova configuração, o fórum público de debate passa a ser exercido através das plataformas, especialmente em redes sociais, pelo que se transfere um grande poder às plataformas digitais, que passam agir com verdadeiros árbitros da vida pública, além de ter sob seu domínio uma constelação de dados pessoais dos usuários, que são tratados e utilizados num capitalismo de vigilância, gerador de muitas discriminações, especialmente pela *modus operandi* de categorização de usuários (Sousa, 2022 p. 53-54).

Como já defendemos em oportunidade anterior[2], nestes ambientes há uma arquitetura permissiva de um maior protagonismo de atores privados, detentores das tecnologias que ditam o ritmo no espaço digital, nomeadamente as grandes plataformas digitais, que passaram a ter tamanho poder econômico, social e político, exercendo, inclusive, a normatização e interferência em direitos fundamentais de utilizadores, especialmente aqueles ligados à liberdade de expressão, autonomia privada, privacidade e proteção de dados pessoais.

Como indicado em tópico anterior, presenciamos na última década um pernicioso modelo de negócios, baseado em utilização de algoritmos e dados pessoais para produzir perfilamento de utilizadores, objetivando direcionar marketing e assuntos de maior interesse, à vista de entreter o usuário o maior tempo possível dentro da plataforma, -gerando maiores lucros- o que acabou por desaguar numa polarização política acentuada, resultando no malbaratamento dos princípios da *rule of law* e a erosão democrática ao redor do mundo. Este ambiente on-line, especialmente dentro das plataformas digitais, passou a ser entendido como a "nova esfera pública" dos tempos modernos.

Resta-nos, de antemão, perceber o que se entende por domínio público ou esfera pública e sua evolução até o presente momento, com especial atenção à transferência de poder que ocorreu da figura do Estado para as empresas de tecnologia, nota marcante deste momento histórico.

Portanto, para IMMANUEL KANT *apud* Morais (2020, p. 45), o Estado deveria ser, necessariamente, dirigido por uma vontade racional, traduzida numa relação de domínio caracterizada pela busca do bem comum, por leis justas e objetivas, acatadas por governante e governados e pela exclusão do arbítrio da violência injustificada, no exercício do poder de autoridade, de forma que deveria pertencer ao próprio Estado o monopólio da violência, justamente como via de estabilidade e paz entre os governados.

As constituições, desde sua origem, são elaboradas com o objetivo de limitar poderes governamentais, protegendo assim os indivíduos da

[2] Cf. FARINHO e MÜLLER DORNELAS, 2024.

interferência de autoridades públicas. Do ponto de vista do direito constitucional, a noção de o poder tem sido tradicionalmente atribuída às autoridades públicas e necessitam de sistemas de freios e contrapesos para dar forma à tão aclamada separação e limitação dos poderes instituídos.

As constituições definem as regras e processos fundamentais de uma comunidade política, e classicamente, o termo refere-se aos mecanismos que delimitam os limites do poder de um estado sobre seus cidadãos. O aspecto substantivo central do pensamento constitucionalista está incorporado nessa necessidade de controlar, limitar e restringir o poder do Estado (Gill; Redeker; Gasser, 2015).

Para o professor catedrático da Faculdade de Direito da Universidade de Lisboa (FDUL), CARLOS BLANCO DE MORAIS (2020, p. 47), a necessidade da criação do Estado vem da ideia de proteger o indivíduo e possibilitar seu crescimento e desenvolvimento pessoal, através do desenho claro de limites de poder do Estado e dos particulares. Logo, é de fácil conclusão que os abusos perpetrados na nova esfera pública digital, notadamente ligados à direitos que estão umbilicalmente atrelados ao crescimento e desenvolvimento da pessoa humana- como os direitos da personalidade-, ferem a própria noção do Estado e, assim, devem ser objeto de atenção e estudo do Direito Constitucional.

Embora formalmente os Estado sejam domínios regidos pelo direito, destaca-se nos dias de hoje, como paradigma de uma evolução civilizacional iniciada pelo movimento constitucionalista do Século XVIII, um modelo de Estado regido pelo primado da Constituição, pela separação dos poderes, pelo princípio da submissão da Administração pública à lei e pela salvaguarda dos direitos dos cidadãos através de Tribunais independentes, ou seja, trata-se do Estado de Direito (Morais, 2020).

BLANCO DE MORAIS (2020, p.47) prossegue afirmando que é o Estado o ente que detêm, de forma mais perfeita, o monopólio do uso da força para fazer cumprir, junto das pessoas, individuais e coletivas, as regras que dele promanam. Esse vínculo de obediência prestada por uma comunidade humana

a uma autoridade que nela impera reconduz-se à ideia weberiana do “Princípio da Dominação”[3].

Implica, pois, a suscetibilidade dos membros de um grupo determinado obedecerem a comandos, gerais e específicos, manifestando um mínimo de vontade de acatar o poder de autoridade de onde brotam os referidos comandos:

> “Uma comunidade onde a autoridade não faça cumprir as suas decisões e onde impere recursivamente a desobediência, desagregar-se-á na anomia, divisão e violência, fazendo os indivíduos valer erraticamente os seus interesses, particulares ou grupais, através da força” (Morais, 2020, p.7).

Todavia, ante a modificação do paradigma social, político, econômico e jurídico, que se nota com o advento da sociedade algorítmica, ocorre a transferência do tradicional domínio da esfera pública para as plataformas digitais, figurando esses atores privados como legítimos detentores de poder econômico[4] - sem precedentes - e pautados por uma visão liberal de internet, onde não existe muito apreço aos limites da Constituição.

Por isso, quando o Estado é cooptado por forças – tanto estatais quanto não-estatais- que antagonizam e desprezam as regras do jogo democrático- como vem ocorrendo na sociedade algorítmica- e passa a existir não mais para os objetivos democráticos, naturalmente há uma tendência à regulação e imposição do Estado de Direito.

Destarte, para o renomado jurista alemão HOFFMANN-RIEM (2021, p.6):

> [...] a transformação digital desenvolveu-se inicialmente com base em estruturas ultrapassadas, incluindo a ordem anterior do Estado, da economia e da sociedade e, por sua vez, encontrou um sistema jurídico que se expandiu no curso do desenvolvimento histórico, como por exemplo o Direito Público nacional, o Direito Civil e o Direito Penal, incapaz de compreender e dar respostas às demandas do ciberespaço.

---

[3] Para Max Weber a dominação é sempre resultado de uma relação social de poder desigual, onde se percebe claramente a existência de um lado que comanda (domina) e outro que obedece, ou seja, há uma relação de subordinação ou verticalidade entre os sujeitos. Os três tipos puros de dominação da dogmática weberiana são: dominação legal, dominação tradicional e dominação carismática. Para maiores aprofundamentos, conferir a obra WEBER, Max. **Textos Coligidos**. São Paulo: Ática, 2001.

[4] A sigla GAMAM é utilizada para designar as cinco maiores *big techs* do mercado que, juntas, somaram em fevereiro de 2023 um valor de mercado de mais de 7 Trilhões de dólares, à saber: Google, Apple, Microsoft, Amazon e Meta

Em vista da globalização dos desenvolvimentos, o direito como um todo é afetado na medida em que o sistema legal contém competências e diretrizes para configurar a ordem social, isso também afeta o agora importante processo da transformação digital e seus resultados. Logo, a proteção pelo direito deverá ser mais abrangente, englobando demais áreas científicas de todos os aspectos da vida social, vez que configuram a nova esfera pública:

> A necessidade de ajudar a moldar futuros desenvolvimentos por meio da lei afeta basicamente todos os usos possíveis das tecnologias digitais. A visão deve ser ampliada tanto em termos sócio-políticos, como jurídicos, ou seja, para incluir as oportunidades e os riscos da digitalização no Estado e na sociedade (Hoffmann-Riem, 2020, p. 7).

Assim, a esfera pública institucionaliza a relação entre indivíduos, democratizando e construindo a opinião pública, bem como assumindo a garantia de deveres e liberdades para os participantes- assentes na liberdade de comunicação dos seus intervenientes e nas consequências jurídicas, previamente delimitadas, de transbordamento dos limites do exercício das liberdades-.

THOMAS VESTING (2021) preconiza que a formação da esfera pública é extremamente dependente das tecnologias disponíveis na sociedade para produzir e fazer circular a informação social:

> A esfera pública liberal, por exemplo, caracterizou-se pela sua emancipação dos contornos centralistas da corte e pela nova estruturas impessoais garantidas pela dinâmica das grandes cidades oitocentistas século. Com o surgimento dos meios de comunicação de massa no século XX, a esfera pública começou a ter um contorno mais pluralista e orientado para o grupo, deixando de consistir essencialmente de indivíduos que estiveram em locais públicos para debates sobre temas gerais. Uma terceira fase, a atual, transformou a esfera pública centrada nos grupos em uma nova constelação gerada pela lógica algorítmica das redes sociais (Vesting, 2021, p. 149).

Por isso, a formação de novos padrões legais para lidar com os efeitos negativos da nova esfera pública digital não pode mais ser guiado pelos padrões anteriores e deve se concentrar em particular em promover a auto-organização do setor tecnológico em questão.

Para que isto funcione bem, JACK BALKIN (2017) advoga que:

> "Os defensores dos princípios do Estado de Direito devem trabalhar em conjunto com as empresas de tecnologia para enfatizar a responsabilidade social destas na nova esfera pública digital, a fim de que percebam e aceitem seu papel central" (Balkin, 2017, p. 65).

Dessarte, VESTIN (2021) ensina que a esfera pública sempre se baseou numa estrutura pluralística de grupo de interesses, porque o pressuposto da opinião é determinado por questões e contribuições de grupos sociais e organizações de interesse, tais como por partidos políticos, por organizações sociais, pelas igrejas, sindicatos, e mídia tradicional.

Lado outro, como bem aponta o pesquisador JOÃO TORNADA (2023, p.6), "[...] a democracia representativa não possui apenas natureza formal, mas também seu viés material, de vontade do povo". Em outras palavras, a realização da vontade do povo não é apenas a representação da vontade da maioria, nem sua correspondência com concepções *a priori* de justiça, mas também o resultado de um processo dialético entre representantes e aqueles eles representam.

Esse modelo discursivo deliberativo é baseado, portanto, em um ideal suposição de comunicação voltada para a obtenção de compromissos e consensos em torno do melhor argumento, legitimando, por via processual, soluções boas, justas e razoáveis.

E, JÜRGEN HABERMAS *apud* TORNADA (2023), conclui que, de acordo com a teoria da ação comunicativa, a lei só seria legítima se fosse produto de um processo deliberativo de discursivo livre e inclusivo, entre os tomadores de decisão e a opinião pública, pelo que não há legitimidade democrática sem uma esfera pública robusta, suficientemente ativa e dialógica para promover, por meio de procedimentos de discursos deliberativos, o escrutínio do poder, o livre desenvolvimento da personalidade dos cidadãos e sua autodeterminação política democrática.

Portanto, as teorias da democracia discursivo-deliberativa estão associadas ao papel da sociedade civil (descentralizada por natureza), como o formador da "opinião pública em um diálogo constante com a classe política

(tradicionalmente institucionalizadas), que ocorre em uma esfera pública aberta, robusta e plural, sendo que a mídia desempenha um papel extremamente importante no acompanhamento e escrutínio da atividade política, contribuindo para a formação da opinião pública e sobre o exercício do poder soberano, bem como na a criação de alternativas (Tornada, 2023).

À evidencia, nas sociedades complexas, a esfera pública “consiste numa estrutura intermediária entre o sistema político e os setores privados, sendo o discurso público fator determinante para robustecer os assuntos políticos (Tornada, 2023).

Logo, o diálogo entre a sociedade civil e o poder institucionalizado depende de liberdades fundamentais assentes na ideia democrática de liberdade de expressão, autonomia privada e uma imprensa livre, crítica e engajada em um “debate público crítico”, pois o Estado, obrigado a responder às questões colocadas pela opinião pública, passaram a utilizá-lo como bússola de moralidade e legalidade para suas ações.

À evidência, o ciberespaço criou um ambiente fértil para uma esfera pública mais ampla, inclusiva, democrática e participativa do debate, em qual o discurso público e privado se combina, transferindo o eixo da tradicional esfera pública analógica, para a novel esfera pública digital.

Soma-se a isto dois fatores imprescindíveis, quais sejam, i) tecnologia estar cada vez mais acessível às pessoas; e ii) a democratização do acesso à informação: assim, na nova esfera pública das plataformas digitais, um sem número de pessoas, com acesso à informação e tecnologias que permitem o compartilhamento de informações, conhecimento e dialogam sobre tudo o que se possa imaginar, especialmente assuntos ligados à coletividade.

Como consequência prejudicial deste ambiente, o excesso de informação -que leva à desinformação- e a manipulação com notícias falsas, percebeu-se uma tendência à polarização de opiniões, aumento do discurso de ódio e intolerância ao contraditório, notadamente em questões políticas.

Nota-se que o modelo de negócio utilizado pelas *big tech*, baseado em dados e algoritmos, especialmente há o *profiling* de utilizadores, a fim de

mantê-los o maior tempo possível online, com sistemas de recomendação e moderação por sistemas de IA, o que contribui sobremaneira para este estado de coisas (Farinho; Müller Dornelas, 2024)[5].

Por isso, o pesquisador SIMÃO SOUSA (2022) faz o alerta que o populismo crescente no mundo fora impulsionado pelas plataformas digitais, que fazem *escola* no preenchimento de todos os espaços da esfera pública digital e são favorecidos por uma campanha de desinformação e polêmicas que, mercê de uma sociedade cada vez mais polarizada e dividida, vão aumentando o grau de influência, inclusive em processos eleitorais democráticos, minando o debate com um discurso populista radical, que apelam para fortes emoções nos utilizadores e influenciam o ciclo vicioso da recomendação algorítmica, gerando lucros ao final.

Assim, THOMAS VESTING (2021) ensina-nos que, no lugar da esfera pública pluralista de grupos de interesse, encontramos, atualmente, uma esfera pública fragmentada em redes amplamente díspares, que são o resultado da grande escala desagregação dos fluxos de comunicação social.

Nesses novos modos de vida, grupos e organizações sociais tradicionais em particulares perdem sua importância e são substituídos por processos muito mais instáveis, como o *locus* das plataformas digitais, que desagua num aumento do individualismo e da intolerância e desincentivo ao contraditório.

Como dito anteriormente, isso se dá em grande parte pelo uso de técnicas adaptativas de algoritmos para impulsionar e recomendar conteúdos que recompensam especificamente postagens que provocam emoções fortes e imediatas interações e que, em suma, criam para o usuário seu próprio mundo singular, levando às consequências indesejáveis, como o conhecido efeito bolha (*bubble effect)* e as câmaras de eco (*echo chamber)*.

Uma das razões que levam à polarização na esfera pública digital, para THOMAS VESTING (2021) é a precisa falta de contraditório, criada por essas

[5] Cf. Farinho, Domingos; Müller Dornelas, Felipe. *Op cit*., 2024 (no prelo).

“bolhas” de pensamento único, o que conduz a um empobrecimento da opinião pública e, por consequência, do debate democrático:

> Há uma autolimitação temática questionável, como se fosse um tipo de “cegueira narcísica” para as realidades da vida fora do próprio grupo, que dificilmente permite mais uma apreciação recíproca das visões de mundo e ideologias de outros meios culturais, de forma a empobrecer o debate democrático e a opinião pública consequentemente (Vestin, 2021, p.159).

E porque as novas “câmaras de eco” são muito menos formalizadas, muito menos institucionalizadas, e em um grau muito menor estruturado pela lei formal do que a mídia da esfera pública pluralista de grupos de interesse, fenômenos totalmente novos como a desinformação e *fake news* tornam-se possíveis.

Corroborando a ideia de uma esfera pública digital marcadamente fragmentada e difusa, SIMÃO SOUSA (2022) constata que este modelo é capaz de proliferar estratégias de desinformação e manipulação algorítmica, contribuindo para radicalização das ideias, dificultando um debate democrático e racional sobre os temas essenciais de uma democracia:

> A definição de esfera pública digital parte de um conjunto amplo de pressupostos que assente desde logo na fragmentação atual do espaço público digital promovido pelo recentrar do debate público com recurso a plataformas digitais, permitindo que a formação da opinião pública seja feita, fundamentalmente, mediante o conteúdo que se encontram em rede, sejam eles verdadeiros ou falsos, proliferando estratégias de desinformação e manipulação algorítmica, contribuindo para radicalização das ideias, dificultando um debate democrático e racional sobre os temas essenciais de uma democracia (Sousa, 2022, p.76).

Assim, a digitalização do mundo global, veio proceder a uma modificação da esfera pública, que passou para o mundo digital proporcionado pelas plataformas, assumindo estas um papel determinante na nova esfera pública digital.

Dado o papel central assumido – propositalmente ou não- na esfera pública digital, as plataformas digitais passam a ser encaradas como verdadeiros sujeitos de direito público e, como tal, estão, sob a égide do Direito Constitucional

e, consequentemente, das formas de limitação de poder e proteção de direitos e garantias fundamentais (Sousa, 2022).

À vista disso, na sociedade algorítmica, as principais ameaças às democracias constitucionais não provêm mais exclusivamente do poder público, uma vez que eles vêm principalmente de atores privados que governam espaços que são espaços formalmente privados, mas exercendo na prática, e sem qualquer salvaguarda, funções tradicionalmente atribuídas às autoridades públicas sem qualquer salvaguarda (De Gregorio, 2022).

Partindo destas premissas, é forçoso reconhecer que a esfera pública digital deva ser tutelada pelo Estado de Direito, ainda que exista grande margem de normatividade própria- como aponta GEOVANNI DE GREGORIO (2023), que denomina tal movimento de "Rule of Tech" -, não se admitindo a mera observação do fenômeno por parte do direito, ante a exasperação do exercício do poder legado às empresas de tecnologia e a forma abusiva que tratam dos direitos dos usuários nesta nova *praça* pública:

> "As tecnologias de inteligência artificial, e particularmente a tomada de decisões automatizada baseada em 'machine learning' e 'deep learning', fornecem outra fonte normativa generativa que molda a 'sociedade algorítmica'. Embora os intervenientes públicos regulem as tecnologias digitais ou os gigantes tecnológicos imponham direitos e liberdade à escala global com base nos seus termos de serviços, a tecnologia também expressa uma forma de governação que escapa à lógica dos intervenientes públicos e privados, desafiando assim os limites tradicionais da 'rule of law', ou mesmo da 'rule of the platform'. Esse entendimento pode ser chamado de 'rule of tech'." (De Gregorio, 2023, p.5, tradução livre). [6]

Consequentemente, é dizer que o ciberespaço deve estar sob os ditames constitucionais, singularmente a limitação de poder das grandes empresas de tecnologia, o respeito aos direitos fundamentais no interior destas e o respeito aos princípios democráticos.

[6] No original: *Artificial intelligence technologies, and particularly automated decision-making based on machine learning and deep learning, provide another generative normative source that shapes the algorithmic society. While public actors regulate digital technologies or tech giants enforce rights and freedom on a global scale based on their terms of services, technology also expresses a form of governance that escapes the logic of public and private actors, thus challenging the traditional boundaries of the rule of law, or even the rule of the platform. This understanding can be called the rule of tech.*

Assim, os poderes públicos ainda desempenham um papel crítico papel na gestão dos espaços digitais e na interferência com direitos e liberdades. No entanto, a influência dos intervenientes privados no ambiente digital suscita cada vez mais preocupações sobre a forma como estas entidades desempenham funções de interesse público ou, em alguns casos, espelham o exercício dos poderes públicos.

Na sociedade algorítmica, portanto, as empresas privadas transnacionais, principalmente as plataformas on-line, exercem poderes governando os espaços digitais, pelo que os conteúdos e os dados coletados podem ser facilmente disseminados à escala global, como se nota das plataformas de redes sociais que implementam tecnologias algorítmicas para moderar e recomendar conteúdo.

Portanto, é uma consequência natural deste novo panorama social a constitucionalização das plataformas, tendo em vista a nova esfera pública digital perpassar por questões que escapam à interesses estritamente privados, adentrando em temas de relevância coletiva que, dada suas influências sistêmicas, devem ser tuteladas pelas normas constitucionais.

1.3) Proteção de dados pessoais e IA no contexto do ciberespaço:

À evidência, a dignidade da pessoa humana, como autonomia individual e controle da reserva de intimidade e apresentação pública na sociedade digital, está no epicentro do direito a proteção de dados pessoais. A fundamentação teórica da proteção de dados pessoais, portanto, está consubstanciada na proteção à autodeterminação informacional - conectada à personalidade - e, por fim, à dignidade da pessoa humana.

Como já exposto em trabalho anteriormente realizado[7], ao lado da proscrição de subjugação[8] e exclusão e, também, da vedação à utilização de

[7] Conf. MÜLLER DORNELAS, 2022.

[8] “A partir da capacidade de prestação e de representação da própria dignidade, não são admissíveis, em princípio, outras interferências estatais na autodeterminação individual acerca do sentido e dos planos que cada um projeta e desenvolve para a própria vida que não sejam estritamente derivadas da necessidade de garantir a reciprocidade do respeito pela igual dignidade de todos e pelo estatuto de humanidade imprescindível da pessoa” (Novais, 2018. p. 115).

coisificação da pessoa -ou fórmula do objeto[9]-, o controle sobre a identidade, a reserva da esfera íntima e apresentação da pessoa integram o conteúdo da dignidade da pessoa humana.

Dessarte, os direitos da personalidade são, por natureza, mais próximos à dignidade da pessoa humana e é nela que mais direta e imediatamente repercutem as exigências de respeito à dignidade humana. Eles se fundam na própria existência do seu titular considerado como *persona*. Neste plano é que a ligação entre direitos fundamentais e a dignidade da pessoa humana toma forma de uma não eventual sobreposição, mas de uma distinção dos respectivos comandos normativos através de seu conteúdo essencial.

Em adição, os direitos fundamentais muitas vezes não são suficientes para tutela efetiva de toda extensão da dignidade da pessoa humana, especialmente diante do surgimento de novas tecnologias e mudanças de paradigmas nas sociedades modernas.

A dignidade humana, todavia, relega um campo residual que não permite uma identificação total com direitos fundamentais, sob pena de ser inócuo e sem conteúdo autônomo. Ou seja, considerados no âmbito normativo, os direitos fundamentais admitem cedência e limitação, enquanto a dignidade humana, como *princípio dos princípios* que sustenta todo arcabouço do Estado de Direito, possui força jurídica imperativa, sendo que esta parcela residual, não alcançada pelos direitos fundamentais, só pode ser objeto de guarida pela dignidade.

É no domínio dos direitos da personalidade, por conseguinte, que existe uma forte aproximação entre o âmbito protetivo dos direitos fundamentais e da dignidade da pessoa humana. Consoante a doutrina do Professor JORGE REIS NOVAIS (2018), existe uma ligação estreita entre direitos da personalidade e a dignidade humana:

[9] "[...] instrumentalização significando que se utiliza alguém apenas como mero meio, com um sentido ou um efeito denegridor, desqualificante. Só no segundo tipo de situação é legítimo falar em eventual violação da dignidade da pessoa humana, ainda que, objectivamente, haja instrumentalização num e no outro caso" (Novais, 2018. p.130).

> Se há direitos que, por natureza, estão mais próximos ou mais intimamente associados à dignidade da pessoa humana, esses são os chamados direitos fundamentais da personalidade, ou seja, aqueles que respeitam e se fundam na própria existência do seu titular considerado como persona, incluindo-se, aí, as garantias jusfundamentais de proteção da vida, da integridade física e psíquica, da liberdade geral de ação e de uma esfera pessoal reservada (Novais, 2018, p. 193).

Por conseguinte, colmatando a lacuna entre a relação da dignidade da pessoa humana com os direitos da personalidade e a proteção de dados, é imperioso regressar, historicamente, ao ano de 1977, pois a Alemanha já possuía uma lei federal de proteção de dados pessoais (*Bundesdatenschutzgesetz* - BDSG), sendo que igualmente, desde 1970, já existira, no *Land* do Hesse, uma lei regional de proteção de dados pessoais, predecessora, portanto, da lei federal.

Todavia, o caso emblemático e amplamente divulgado como um marco teórico do reconhecimento da proteção de dados pessoais se deu em 1982, através da lei do censo alemão, onde tal continha previsão de uma consulta à população, com mais de cento e sessenta perguntas, que, em seguida, fora digitalizada, sistematizada e inserida em bancos de dados da administração, levantando diversos questionamentos sobre a vigilância e a falta privacidade[10].

Nesse sentido, acabou por ser impugnada no Tribunal Constitucional Alemão (*Bundesverfassungsgericht- BvG*), que declarou a inconstitucionalidade desta lei, nomeadamente pela possibilidade de vigilância estatal dada a ausência de finalidade do tratamento, reconhecendo um direito à autodeterminação informacional (*Informationelle Selbstbestimmungsrecht)* como decorrência direta da dignidade da pessoa humana.

A partir deste célebre julgamento pelo *BvG*, iniciou-se uma preocupação maior no estudo sistemático da proteção de dados, que resultou anos depois em numa tentativa de estabelecimento robusto de um quadro-legal sobre o tema. Por isso, a proteção de dados pessoais foi reconhecida como um direito fundamental pela Carta de Direitos Fundamentais da União Europeia,

[10] Cf. MÜLLER DORNELAS, Felipe. *Op. Cit*., 2022

através do artigo 8º, n.º 1[11], bem como pelo artigo 16º n.º 1[12], do Tratado sobre o Funcionamento da União Europeia (TFUE).

Igualmente, antes da entrada em vigor do Regulamento (EU) 2016/679, denominado de Regulamento Geral de Proteção de Dados (RGPD), a proteção de dados pessoais já era reconhecida através da Directiva 95/46/CE, posteriormente revogada em prol da utilização de um instrumento normativo dotado de vinculação obrigatória direta aos Estados-Membros- como é o caso do instituto jurídico do Regulamento- e capaz de uniformizar conceitos e procedimentos em matéria de dados pessoais em todo o espaço europeu[13].

Somado a isso, a sociedade vem experimentado um exponencial crescimento das tecnologias e a consequente utilização massiva de dados pessoais, que funciona como impulsionador deste novo sistema econômico e social[14].

Recorrendo novamente aos relatos da socióloga SOSHANA ZUBOFF (2020), é patente a vigilância massiva exercida sobre as pessoas nos tempos atuais, notadamente por meio das grandes empresas de tecnologia, destacando que os dados pessoais funcionam como verdadeiras *comodities* dos "novos tempos".

Nesse sentido, é o próprio considerando n.º6 do Regulamento Geral de Proteção de Dados (RGPD), que assim determina:

> A rápida evolução tecnológica e a globalização criaram novos desafios em matéria de proteção de dados pessoais. A recolha e a partilha de dados pessoais registaram um aumento significativo. As novas tecnologias permitem às empresas privadas e às entidades públicas a utilização de dados pessoais numa escala sem precedentes no exercício das suas atividades. As pessoas singulares disponibilizam cada vez mais as suas

---

[11] *Todas as pessoas têm direito à protecção dos dados de carácter pessoal que lhes digam respeito.*

[12] *Todas as pessoas têm direito à proteção dos dados de caráter pessoal que lhes digam respeito.*

[13] Precisamente neste sentido é o considerando n.º10 do Regulamento que assim dispõe: *A fim de assegurar um nível de proteção coerente e elevado das pessoas singulares e eliminar os obstáculos à circulação de dados pessoais na União, o nível de proteção dos direitos e liberdades das pessoas singulares relativamente ao tratamento desses dados deverá ser equivalente em todos os Estados-Membros. É conveniente assegurar em toda a União a aplicação coerente e homogénea das regras de defesa dos direitos e das liberdades fundamentais das pessoas singulares no que diz respeito ao tratamento de dados pessoais. [...]*

[14] Ficou célebre a frase do matemático britânico Clive Humby quando disse que "Data is the new oil".

> informações pessoais de uma forma pública e global. As novas tecnologias transformaram a economia e a vida social e deverão contribuir para facilitar a livre circulação de dados pessoais na União e a sua transferência para países terceiros e organizações internacionais, assegurando simultaneamente um elevado nível de proteção dos dados pessoais (RGPD)[15].

Portanto, a proteção dos dados pessoais[16] ocupa uma centralidade amplamente reconhecida na esfera pública digital, sendo uma das preocupações do domínio público a sua devida tutela, especialmente por se tratar do direito em essência que faz fluir este novel arranjo econômico, como igualmente reconhecido pela agenda digital europeia, tanto como diretriz, como direito a ser tutelado.

Nas palavras do pesquisador BRUNO BIONI (2021), esta nova forma de organização da sociedade, baseada na tecnologia, tem a sua centralidade consubstanciada na coleta para tratamento dos dados pessoais dos indivíduos, visando à transformação destes em informação que, por conseguinte, impulsiona todo um sistema econômico:

> [...] no estágio atual, a sociedade está encravada por uma nova forma de organização em que a informação é o elemento nuclear para o desenvolvimento da economia, substituindo os recursos que outrora estruturavam as sociedades agrícola, industrial e pós-industrial (Bioni, 2021, p. 5).

Portanto, a evolução tecnológica proporcionou o tratamento de uma gama infindável de dados, em alta velocidade, modificando a estrutura das relações sociais para uma nova era, em que se encurtam as distâncias de tempo e espaço e, consequentemente, as relações políticas, econômicas, sociais e culturais tomam nova forma.

Destarte, com a peculiar precisão, PAESANI (2010), assevera que houve um "encolhimento" do mundo, por meio da compreensão do espaço-tempo, isto quer dizer, a disrupção causada pelo avanço tecnológico,

---

[15] Disponível em https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=celex%3A32016R0679. Acesso em 06 de julho de 2023.

[16] Dados Pessoais, segundo a conceituação legal trazida no número 1) do artigo 4.º da RGPD, é toda informação relativa a uma pessoa singular identificada ou identificável («titular dos dados»); é considerada identificável uma pessoa singular que possa ser identificada, direta ou indiretamente, em especial por referência a um identificador, como por exemplo um nome, um número de identificação, dados de localização, identificadores por via eletrónica ou a um ou mais elementos específicos da identidade física, fisiológica, genética, mental, económica, cultural ou social dessa pessoa singular;

especialmente a internet, proporcionou uma escalabilidade das interações sociais, numa velocidade jamais vista, desaguando numa nova percepção de mundo, onde a velocidade da informação é praticamente instantânea, desconhecendo limites territoriais ao redor do globo.

Este contexto, também possibilitou uma vigilância massiva exercida por Estados, governos e empresas, por meio dos dados pessoais de cidadãos e usuários, permitindo a categorização das pessoas em verdadeiros "perfis virtuais", que passam a fundamentar a tomada de decisões comerciais, políticas e afetivas e, por consequência, surgiram uma gama de problemas oriundos deste *modus operandi*.

BRUNO BIONI (2021) exemplifica concretamente este estado de coisas, quando nos diz que as informações sobre os hábitos de consumo dos cidadãos permitem empreender de forma mais eficiente no mercado, aumentando as possibilidades de êxito, melhorando a segmentação de um produto ou serviço.

Ou seja, a informação convertida em conhecimento é tida como matéria-prima de uma economia digital -ou do capitalismo de vigilância-, que tem, através dos dados pessoais dos cidadãos, seu "combustível" operacional.

Diante desse cenário, os direitos fundamentais das pessoas são constantemente violados, seja por vigilância, seja por tratamento inadequado e ilegal de seus dados pessoais, seja por manipulação direta e indireta a partir de uma categorização e "estereotipação" do usuário tem causado enormes danos não apenas a estes, mas também ao próprio Estado de Direito, em razão dos efeitos sistêmicos destas violações.

Assim, é importante destacar que, neste contexto, a tecnologia permitiu *escalar* a organização e sistematização de dados pessoais com o surgimento da inteligência artificial potencializando os efeitos nefastos deste quadro social, político e econômico.

Partindo da premissa de que os dados pessoais representam o estado primitivo da informação, é natural de se esperar que as referidas bases devam ser processadas e organizadas para se transformarem em inteligência e, por meio disso, ganhar utilidade. Inicialmente, essa sistematização e esse

processamento de dados utilizavam uma lógica de entrada (*input*) e saída (*output*), com uma necessária gestão manual ou automatizada de um grande conjunto de elementos, conforme nos ensina o saudoso doutrinador DANILO DONEDA (2019).

Conforme ensina-nos LAURA SCHERTEL MENDES (2021. p. 424), a tecnologia permitiu um salto qualitativo e quantitativo no processo de transformação de dados pessoais (matéria-prima bruta) em informação, primeiramente com utilização de *softwares* e, posteriormente, por meio da inteligência artificial e *big data*[17].

Recorrendo mais uma vez ao polido ensinamento de BRUNO BIONI (2021), o *big data* representa o *êxtase* do processo de gestão e sistematização dos dados em informação, por meio da velocidade, do volume e da variedade que esses sistemas são capazes de processar:

> Por isso os dados passaram a ser analisados não mais em pequenas quantidades ou por amostras, mas em toda a sua extensão. Há um salto quanto ao volume de dados processados, tornando possível correlacionar uma série de fatos (dados), estabelecendo-se entre eles relações para desvendar padrões e, por conseguinte, inferir, inclusive, probabilidades de acontecimentos futuros (Bioni, 2021, p. 37).

Ocorre que, ao promover esse *salto* relacionado ao volume, à variedade e à velocidade (3 V's) de dados tratados nas bases, transformando-os em informação, criando-se padrões de perfil humano e encaixando as pessoas em categorias, a partir desses padrões *identificados* pela máquina, surgiu um campo fértil e propício para violação de direitos fundamentais, especialmente o princípio da igualdade, por meio de *perfilizações* discriminatórias.

[17] Outrossim, para Mayer-Schönberger e Cukier, não há definição precisa de *Big Data*, mas o fenômeno pode ser caracterizado por três tendências. Em primeiro lugar, a quantidade de dados e informação coletada. As análises de *Big Data* não apenas reúnem mais dados do que nunca, mas buscam juntar todos os dados e informações referentes a uma situação particular não somente uma amostra deles, como colocam os autores, em *Big Data*, 'n=tudo'. Em segundo lugar, devido à grande quantidade de informações disponíveis, os dados podem ser imprecisos. Na medida em que a magnitude aumenta, do mesmo modo elevam-se as chances de equívocos. A terceira propriedade é a de buscar correlações, em vez de causalidades. Isso significa que a relação entre dois fatos ou características é determinada de acordo com uma análise estatística".

Chegou-se a ponto de, a partir dos *likes* ou “gostos” em uma rede social, detectar o perfil exato de preferências dos usuários de determinada plataforma *on-line*, identificando com precisão — independentemente da autorização e consentimento do utilizador/consumidor — a porcentagem de usuários homossexuais, heterossexuais, brancos e negros, republicanos e democratas (Bioni, 2021, p. 84).

Para além, é possível identificar que o *big data* se utiliza, em suas análises, de correlações (probabilidade de um evento ocorrer, caso outro evento também se realize) e não de causalidade (agente que liga dois processos, sendo um a causa e outro o efeito) e, por essa razão, poderá apontar situações discriminatórias, por usar essa relação estatística para gerar a informação, ao invés de tentar compreender os eventos a partir da lógica de causa e efeito.

Destarte, precisas são as lições de MAYSON *apud* Lordelo (2022, p. 178), que já identificara a perpetuação de desigualdades enraizadas ou estruturais nesses tipos de mecanismos, baseados em frios cálculos realizados por algoritmos “Em um mundo racialmente estratificado, qualquer método de predição projetará essas desigualdades do passado para o futuro”.

É evidente que o uso desregulado das tecnologias algorítmicas de inteligência artificial, ancorados em plataformas digitais de *big techs*, detentoras de grandes poderes numa sociedade digitalizada, acabam por propiciar vieses cognitivos, preconceitos enraizados e estruturais e opacidade em decisões, como podemos exemplificar por meio do uso de algoritmos e os vieses de gênero, raça, opção sexual e social para reconhecimento facial, escolha de candidatos para ofertas de emprego, oferta de crédito, etc.

Segundo os dados do portal *Social Media Users* 2023, existem 4.9 bilhões de usuários de mídias sociais no mundo, o que representa 60,49% da população mundial conectada às plataformas digitais, com estimativa de que em 2027 esse número suba para 5.85 bilhões de utilizadores. Dessas plataformas, o Facebook é que possui maior número de usuários, com aproximadamente 3.03

bilhões, sendo que os países que mais possuem pessoas conectadas às redes são China, Índia e Estados Unidos da América, respectivamente[18].

Lado outro, o avanço exponencial da Inteligência Artificial (IA) nos últimos anos representa ao mesmo tempo um fator de avanço na sociedade, notadamente pela capacidade viabilizar novos serviços, produtos e tecnologias digitais, mas igualmente uma ameaça à diversos direitos fundamentais, dada a frieza dos algoritmos e sua capacidade de chegar a resultados conflitantes com direitos e geradores de discriminações, vigilância e, principalmente a utilização inadequada/ilícita de dados pessoais.

Em vista disto, o prestigiado pesquisador e professor alemão PHILIPP HACKER (2023a) aponta precisamente que, junto com a temática da mudança climática, a inteligência artificial é o tópico de maior preocupação da atual presidência da União Europeia, nomeadamente pela importância desta tecnologia na vida cotidiana da humanidade – com seus prós e contras-, levando o legislador europeu a empenhar esforços na compreensão e regulação eficaz[19].

Segundo POLLICINO E DE GREGORIO (2021), a inteligência artificial, pode fornecer melhores sistemas de aplicação de regras legais ou melhorar o desempenho dos serviços públicos. No entanto, o domínio referente ao "buraco negro" dos algoritmos, aquele que não se consegue escrutinar - que caracterizam a sociedade contemporânea-, desafiam a proteção de direitos fundamentais e valores democráticos, incentivando ao mesmo tempo os legisladores a encontrar um quadro regulamentar que equilibre o risco e a inovação, considerando o papel e responsabilidades dos atores privados na sociedade algorítmica.

[18] Cfr. https://www.demandsage.com/social-media-users/, Acesso em 23 de outubro de 2023.
[19]Segundo HACKER: *When the decisive breakthrough for the training of deep neuronal networks was achieved in 2006, it was hardly foreseeable that Artificial Intelligence (AI) techniques would find their way into the most diverse areas of economic and private activity within a few years: from credit scoring to personnel recruitment, from autonomous driving to medical diagnostics and research. The European Commission has taken due heed of these developments. Alongside climate change, AI is set to become the second major topic of Ursula von der Leyen's presidency. With the White Paper on Artificial Intelligence, the Commission has now presented a much anticipated blueprint for the promotion, but also for the regulation of this technology.* Disponível em https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3556532. Acesso em 11 de julho de 2023.

Os desafios levantados pelas tecnologias de inteligência artificial não se limitam a liberdade de expressão, privacidade e proteção de dados. As democracias constitucionais estão sob pressão para garantir segurança jurídica e previsibilidade de processos automatizados. processos de tomada de decisão que podem afetar coletivamente os valores democráticos

A IA é uma família de tecnologias em rápida evolução capaz de oferecer um vasto conjunto de benefícios económicos e sociais a todo o leque de indústrias e atividades sociais.

Ao melhorar as previsões, otimizar as operações e a afetação de recursos e personalizar o fornecimento dos serviços, a utilização da inteligência artificial pode contribuir para resultados benéficos para a sociedade e o ambiente e conceder vantagens competitivas às empresas e à economia europeia. Essa ação torna-se especialmente necessária em setores de elevado impacto, incluindo os domínios das alterações climáticas, do ambiente e da saúde, do setor público, das finanças, da mobilidade, dos assuntos internos e da agricultura. Contudo, os mesmos elementos e técnicas que produzem os benefícios socioeconómicos da IA também podem trazer novos riscos ou consequências negativas para os cidadãos e a sociedade.

Por isso, o objetivo da regulamentação são de quatro vieses, quais sejam: i) garantir que os sistemas de IA colocados no mercado da União e utilizados sejam seguros e respeitem a legislação em vigor em matéria de direitos fundamentais e valores da União; ii) garantir a segurança jurídica para facilitar os investimentos e a inovação no domínio da IA; iii) melhorar a governação e a aplicação efetiva da legislação em vigor em matéria de direitos fundamentais e dos requisitos de segurança aplicáveis aos sistemas de IA; iv) facilitar o desenvolvimento de um mercado único para as aplicações de IA legítimas, seguras e de confiança e evitar a fragmentação do mercado.

Em suma, trata-se de tutelar a cibersegurança, a segurança jurídica, *governance* e *enforcement*, inovação e o mercado, todos sempre em linha com os direitos fundamentais e demais valores da União. Extrai-se, pois, que a preocupação maior da União Europeia na tutela desta esfera pública digital, refletida na proposta de regulação dos domínios da inteligência artificial, seria

compatibilizar a promoção e desenvolvimento da IA (*ecosystem of excellence*) com uma regulação protetora de direitos fundamentais (*ecosystem of trus*t).

Nesse sentido, a regulação das tecnologias de inteligência artificial deve centrar-se no ser humano, de modo que as pessoas possam confiar que a tecnologia é utilizada de uma forma segura e em cumprimento da lei, incluindo em matéria de respeito dos direitos fundamentais.

Dessarte, a abordagem antropocêntrica está em linha com a agenda para uma Europa digital até 2030, que expressamente elenca como diretriz o antropocentrismo, onde as tecnologias digitais devem proteger os direitos das pessoas, apoiar a democracia e assegurar que todos os intervenientes digitais agem de forma responsável e segura.

Como mencionado, os principais direitos fundamentais afetados pelo mal uso da inteligência artificial, mormente aqueles ligados aos riscos específicos como imprevisibilidade, autonomia técnica, opacidade, falta de dados técnicos para treinamento e manipulação, acabam, por consequência, atingindo mais fortemente a proteção de dados pessoais, a privacidade e a igualdade, no viés da discriminação levada à efeito pelos algoritmos (Hacker, 2020, p. 5).

Vale trazer à tona a precisa intervenção do professor CHRISTIAN TRONCOSO junto à comissão de juristas nomeados pelo Senado Federal brasileiro (SF) para análise do marco regulatório da Inteligência Artificial no Brasil, que reconhece uma clara interseção entre IA e dados e que caberá à sociedade traçar os limites da regulação, tendo em vista os benefícios e prejuízos do avanço da tecnologia e o respeito aos direitos fundamentais – como a proteção de dados:

> Então, sim, há uma interseção da proteção de dados e IA que levanta várias discussões, mas devemos reconhecer que inovação é importante, mas a proteção de dados também. Nós temos que estar conscientes dessa interseção e decidir, como uma sociedade, onde queremos delinear esses limiares, mas é importante, principalmente, que tenhamos leis transparentes e claras sobre como implementar essas regulamentações e que as empresas tenham orientações claras sobre o que é permitido e onde estão as limitações (Senado Federal, 2022, p. 132).

Igualmente, a Agência Nacional de Proteção de Dados brasileira (ANPD) constatou e explicitou seu entendimento, via “Análise Preliminar” do PL 2338/2023, que regulamenta a Inteligência Artificial no Brasil, de que o projeto legislativo possui diversos pontos de interação com a Lei Geral de Proteção de Dados (LGPD), notadamente no que diz respeito à: a) tutela de direitos; b) à classificação de sistemas de IA de alto-risco e; c) aos mecanismos de governança.

Aduz ainda no parecer que a ANPD apoia o fomento à inovação, como a proposta de criação de ambientes de regulação experimental, como os *sandboxes* regulatórios, desde que sejam desenhados com o propósito de promover a inovação responsável.

Por inovação responsável podemos entender, regressando ao que fora dito pelo professor CHRISTIAN TRONCOSO, no sentido que a ANDP também está preocupada com a regulação, mas aquela que não estrangule o desenvolvimento da tecnologia, compatibilizando os direitos fundamentais em rota de colisão: proteção de dados e liberdade de inovação.

1.4) Direções e endereçamentos possíveis no constitucionalismo digital:

Diante deste panorama é fácil constatar que o estado de direito, os direitos fundamentais e os valores democráticos estão sob pressão na sociedade digitalizada e estas ameaças às democracias constitucionais podem levar ao questionamento sobre o papel da regulação e da política no âmbito da sociedade algorítmica e das tecnologias. Todavia, o debate sobre a regulação das tecnologias digitais perpassa por questões de difícil resolução, como àquelas as noções consolidadas de soberania e território; nos limites da mediação e ponderação de direitos fundamentais entre particulares; na vigilância e proteção de dados; na velocidade da absorção pelo direito diante das novas tecnologias e na própria crença do espírito democrático.

Precisamente, o embate inicia-se no ponto que a regulação, com base em limites geográficos, é inviável, de modo que a aplicação das leis nacionais à Internet é inviável e inócua. DAVID JOHNSON e DAVID POST *apud* ORESTE POLLICINO e GIOVANNI DE GREGORIO (2019) apontaram a problemática da não

obediência das fronteiras pelas tecnologias, em relação a configuração de Estado tradicional – e a respectiva submissão às leis-. Vejamos:

> Eventos na Internet ocorrem em todos os lugares, mas em nenhum lugar em particular e, portanto, nenhuma jurisdição tem uma reivindicação mais convincente do que qualquer outra de submeter os eventos exclusivamente às suas leis (Pollicino; De Gregorio, 2019).

Ou seja, a partir de uma visão *ciber-anárquica*, a ascensão de uma suposta "lei da Internet" tem o condão de causar a ruptura da soberania do Estado para o ciberespaço, tornando qualquer tentativa regulatória irrelevante para o ambiente digital. Isso mostra-se problemático para o a *Rule of Law*, uma vez que a autorregulação do o ciberespaço teria marginalizado as normas jurídicas, criando uma *Rule of Tech*, ou seja, uma normatividade tecnológica própria, como já apontado por DE GREGORIO (2023).

Entretanto, o esforço conjunto de Estados, organizações e sociedade civil e plataformas vem comprovando a possibilidade de um ambiente regulatório, levando em conjunto a criação de quadros-legais típicos; investimentos em literacia digital; promoção de boas práticas de respeito à direitos fundamentais no ambiente tecnológico; assinatura de tratados e convenções para unificação de regras; dentre outras iniciativas de *hard power* em conjunto com *soft power,* que delineiam o caminho para os tipos de regulação possíveis, como a tradicionalmente feita via Estados, e as novos tipos, à teor da autorregulação – de cunho eminentemente privado e autorregulação regulada – realizada em conjunto entre particular e Estado.

Já SIMÃO SOUSA (2022) entende, ao seu turno, que o esforço para a constitucionalização das plataformas digitais deve ser tomado em direção a uma *Governance Digital Multinível,* com o estabelecimento de mecanismos e princípios comuns, compatíveis e atualizados ao espaço digital, partindo de uma regulação em nível internacional, comunitário, doméstico e, nomeadamente, a autorregulação pelas empresas que atuam no digital.

Destarte, para o suprarreferido autor, se é fato que constatamos o surgimento da nova esfera pública digital, é forçoso reconhecer que os serviços prestados pelas plataformas digitais devem ser considerados como um serviço

de interesse econômico geral e, assim sendo, regular-se-á o ambiente digital sob a sujeições e limitações de ordem constitucional:

> Com a imposição de limites decorrentes de uma lógica de governação multinível, proceder-se-á a uma alteração na arquitetura das próprias plataformas digitais e da escolha do que aparece exposto ao utilizador, permitindo salvaguardar os processos democráticos pela limitação e maior controlo de disseminação de informação falsa e dedicada a grupos específicos, assentes em determinado perfis sociológicos, promovendo uma redução da personalização da informação e uma maior abertura à diversidade e pluralidade do conteúdo e informação apresentados, permitindo uma maior fragmentação de conteúdo e melhoria de uma experiência saudável e partilhada (Sousa, 2022, p. 127).

Nessa linha de raciocínio, complementa EDOARDO CELESTE (2019) afirmando que:

> O endereçamento da resposta do direito constitucional aos problemas oriundos das tecnologias digitais, perpassa por um processo multinível de produção normativa, que englobe duas dimensões normativas, não apenas uma dimensão tradicional de constitucionalismo do Estado-nação — fulcrada no âmbito num quadro legal legislativo -, mas igualmente numa dimensão transnacional, baseada em instrumentos advindos tanto de atores estatais, quanto de particulares. (Celeste, 2019, p. 18).

Portanto, arremata SIMÃO SOUSA (2022) que, à vista dos desafios desta nova esfera pública digital, um sistema que contemple uma proteção integral e eficiente do usuário em relação às plataformas deverá ser construído a partir de legislações internas, comunitárias e internacionais, em camadas multiníveis de proteção, para que seja respeitado o pilar fundamental do sistema jurídico pátrio, qual seja, o princípio dignidade humana:

> Como bem se compreenderá, a solução para os problemas que derivam das plataformas, sejam eles para o exercício digital de liberdades constitucionais, seja para a possível restrição da liberdade de expressão na esfera digital perpetrada por um ente privado, sempre residirá num sistema compósito que parta de medidas nacionais em simbiose com medidas europeias e abrangentes que legitimem o poder do cidadão e respeitem o princípio da dignidade da pessoa humana, desumanizada sempre que as decisões que lhe digam respeito sejam tomadas por um algoritmo, mais ou menos evoluído, sendo certo que o cumprimento de obrigações constitucionais para com os utilizadores sempre decorram do princípio da dignidade da pessoa humana. (Sousa, 2022, p. 123)

Podemos utilizar como exemplificação de um *governance multinível*, típico de uma regulação em que se prestigia o constitucionalismo digital, o artigo 14.º, n. 4.) do Regulamento dos Serviços Digitais (RSD), em que determina a observância nos termos e condições das plataformas, no âmbito da União Europeia, dos direitos fundamentais elencados na Carta de Direitos Fundamentais da União Europeia. Vejamos:

> Os prestadores de serviços intermediários agem de forma diligente, objetiva e proporcionada na aplicação e execução das restrições referidas no n.º 1, tendo devidamente em conta os direitos e interesses legítimos de todas as partes envolvidas, incluindo os direitos fundamentais dos destinatários do serviço, como a liberdade de expressão, a liberdade e o pluralismo dos meios de comunicação social e outros direitos e liberdades fundamentais, tal como consagrados na Carta.

Neste tocante, há uma clara determinação do RSD para que exista um diálogo entre a Carta de Direitos Fundamentais da UE e a aplicação de direitos fundamentais pelo particular, na relação privada travada entre usuário e plataforma, a fim de que se prestigiem os direitos e interesses legítimos de todas as partes envolvidas, incluindo os direitos fundamentais dos destinatários do serviço, como a liberdade de expressão, a liberdade e o pluralismo dos meios de comunicação social e outros direitos e liberdades fundamentais

Destarte, exemplificamos também levando em consideração as questões de recomendação de conteúdo, extremamente cara ao debate por envolver o núcleo do sistema do capitalismo de vigilância, pela utilização de dados para lucro, bem como o direito à proteção de dados e IA.

Nesta linha de pensamento, FARINHO; MÜLLER DORNELAS (2024) alertaram para uma série de regras no RSD, em prol daquilo que referia anteriormente SIMÃO SOUSA, ou seja, normas que intentam melhorar a recomendação de conteúdo, em prol de uma constitucionalização das plataformas.

Observemos, pois:

> A comunicação com o ponto único de contacto (cfr. artigo 12.º)[20]; As notificações aos prestadores de serviços de alojamento virtual por parte dos utilizadores que entendem ter encontrado conteúdo em violação dos termos e condições ou de qualquer disposição legal (cfr. artigo 16.º/6); Moderação de conteúdos através da "tomada de decisões algorítmicas" (cfr. artigos 14.º e 17.º/3/c))[21]; [...] No âmbito da atividade Plataformas Online de Muito Grande Dimensão (POMGD) e os Motores de Busca de Muito Grande Dimensão (MBMGD): O RSD considera que a própria forma como a "conceção dos sistemas algorítmicos utilizados"[22] é feita pode criar *riscos sistémicos* (Farinho; Müller Dornelas, 2024, p.19).

Nas palavras de GIOVANNI DE GREGORIO (2022), a resposta aos desafios enfrentados pelo constitucionalismo digital não perpassa meramente por perceber se as democracias constitucionais poderiam introjetar valores democráticos na arquitetura tecnológica, vez que a tecnologia é apenas um meio de mediar a relação de poder entre os humanos. Ou seja, por trás das tecnologias digitais, incluindo a inteligência artificial, existem *players* que definem as características destes sistemas e aí que deve ser concentrado esforços em prol de um ciberespaço devidamente concertado com a normatividade constitucional.

> Portanto, o principal desafio do direito constitucional na sociedade algorítmica não é regular a tecnologia, mas enfrentar as ameaças provenientes da ascensão de poderes privados transnacionais sem responsabilidade, cujos efeitos globais produzem cada vez mais desafios locais para as democracias constitucionais. Num certo sentido, a missão do constitucionalismo moderno é proteger os direitos fundamentais e, ao mesmo tempo, limitar o surgimento de poderes fora de qualquer controle (De Gregorio, 2022, p. 21, tradução livre).[23]

[20] Cfr. Considerando 43.
[21] Cfr. Considerando 45.
[22] Cfr. Considerando 81; cfr. também os Considerandos 84, 85 e 88.
[23] No original: *Therefore, the primary challenge for constitutional law in the algorithmic society is not to regulate technology but to address the threats coming from the rise of unaccountable transnational private powers, whose global effects increasingly produce local challenges for constitutional democracies. In a sense, the mission of modern constitutionalism is to protect fundamental rights while limiting the emergence of powers outside any control.*

Por isso, GIOVANNI DE GREGORIO (2022) segue ensinando que o acumulo de poder pelas empresas de tecnologia ao longo de décadas operando o mercado digital sem regulação – ou mínima-, levou ao acumulo de poder inigualável destes entes privado e, por isso, quando as liberdades se transformam em formas de poder, garantir a supervisão e as salvaguardas democráticas pode impedir que a dinâmica do mercado impulsione os valores constitucionais é algo imperioso e urgente.

> As plataformas on-line têm de tornar-se mais influente operando à sombra dos governos. Desenvolveram as suas funções como procuradores ou entidades delegadas de autoridades públicas para aplicar políticas públicas on-line e de forma autônoma dependem da combinação entre poder de mercado e assimetria tecnológica. O problema do poder privado não é apenas econômico, mas também político. A acumulação de autoridade arbitrária no mercado, fora de qualquer forma de responsabilidade política, pode ser considerada um exercício de poder semelhante que caracteriza o exercício da autoridade pública. Quando as liberdades se transformam em formas de poder, garantir a supervisão e as salvaguardas democráticas pode impedir que a dinâmica do mercado impulsione os valores constitucionais (De Gregorio, 2022, p.30, tradução nossa).

Outrossim, uma solução bem apontada para os problemas enfrentados pelo constitucionalismo digital é aquela trazida por JOÃO LORDELO (2022), qual seja, seria através da utilização da garantia fundamental do devido processo legal digital ou *data due process of law* contra as plataformas, como forma de uma garantia mínima de um ecossistema digital compatível com a ordem constitucional na sociedade digitalizada[24].

Ou seja, o devido processo legal digital funcionaria como uma garantia fundamental de elevada importância no ciberespaço, assegurando a correta aplicação e interpretação dos direitos fundamentais dos usuários afetados ao abrir o contraditório, ampla defesa e recursos, ofertando racionalidade argumentativa às decisões tomadas pelas plataformas e,

[24] Para tanto, o autor enumera alguns princípios que guardam compatibilidade no espaço digital e se mostram capazes de tutelar minimamente um devido processo digital, tais como: i) princípio da auditabilidade; ii) princípio da transparência e direito de explicações contrafactuais; iii) princípio da consistência ou regularidade procedimental; iv) princípio do controle social e v) princípio da precaução.

consequentemente, aferição de compatibilidade aos parâmetros regulatórios vigentes.

Para ilustrar, o Senado Federal do Brasil, através do Projeto de Lei n.º 592, de 2023, pretende deixar expressamente claro na ordem jurídica brasileira a observância obrigatória da aplicação do direito ao contraditório, ampla defesa e recurso nas hipóteses de moderação de conteúdo feita por plataformas digitais.[25]

Em resumo, o devido processo legal digital também é vital na busca de igualdade entre usuários e plataformas digitais, nomeadamente no combate à discriminação, figurando como garantia fundamental de um constitucionalismo digital.

De mais a mais, a tentativa de regulação do novo espaço público deve ser configurada na construção de uma ordem em fragmentos sociais, em particular por estar vinculado a processos de normatividade nos novos campos da alta tecnologia. Uma das tarefas centrais do pensamento jurídico atual é refletir sobre os modelos de governança compatíveis com a nova indústria. Em primeiro lugar, é preciso perceber as novas estruturas sociais que se concentram cada vez mais no desenvolvimento cognitivo e tecnológico bases e moldar a nova indústria dependente de algoritmos. Começando do entendimento dessa nova infraestrutura, faz o segundo momento para abordar medidas regulatórias algo mais plausível com a realidade atual.

Isso demonstra a complexidade de uma regulação temática, nomeadamente pela fragmentariedade da internet e a transversalidade dos envolvidos e dos assuntos, somado à velocidade das mudanças, que o direito deverá enfrentar para tutelar os interesses no espaço digital, pelo que como se revelará ao longo desta dissertação, em consonância com o que reza o constitucionalismo digital, a resposta deverá indicar a participação de todos os envolvidos, ou seja, as empresas, a sociedade civil, os usuários e o poder

[25] “Dos direitos e das garantias dos usuários de redes sociais: Art. 8º-A - Aos usuários, nas relações com os provedores de redes sociais, são assegurados os seguintes direitos, sem prejuízo do disposto na Seção I deste Capítulo: II - contraditório, ampla defesa e recurso, a serem obrigatoriamente observados nas hipóteses de moderação de conteúdo, devendo o provedor de redes sociais oferecer, no mínimo, um canal eletrônico de comunicação dedicado ao exercício desses direitos”.

público, tanto em âmbito nacional, como internacional, em verdadeira sinergia e coordenação de esforços na busca de parâmetros comuns que satisfaçam os limites da *rule of law* das democracias maduras.

2) Inteligência Artificial (IA):

2.1) Conceitos e panorama histórico:

Dados são a mola-mestra da sociedade atual baseada em sistemas computacionais altamente tecnológicos. Na *sociedade algorítmica*, a normatividade originada das regras algorítmicas (*rule of tech*) estão se tornando cada vez mais importantes, em vista da utilização destes sistemas de ordenação para resolução de problemas, permeando um emaranhado complexo de arranjos sociais e econômicos, como podemos vislumbrar a partir das plataformas digitais, que influenciam sobremaneira todos os campos da vida cotidiana, erigindo graves e importantes questionamentos de ordem ético-jurídicas.

Dessarte, para SHOSHANA ZUBOFF (2020) o *Capitalismo da Vigilância* migrou da mera coleta de dados pessoais de usuários -que ocorria principalmente por meio de plataformas digitais, especialmente redes sociais-, para um *business* que compreende a negociação de dados coletados, com fito de uma sistematização preditiva dos referidos dados, com larga utilização de sistemas de inteligência artificial, capazes de antever comportamentos humanos e, dessa forma, induzir os usuários a consumos personalizados, o que elucida o desrespeito sistêmico à direitos fundamentais no ciberespaço.

À semelhança de muitas outras tecnologias, os sistemas de IA tanto podem favorecer como prejudicar a fruição dos direitos fundamentais. Podem beneficiar as pessoas, por exemplo, ajudando-as a rastrear os seus dados pessoais ou aumentando o seu acesso à educação e apoiando, assim, o direito à mesma. No entanto, dado o alcance e a capacidade dos sistemas de IA, também podem afetar negativamente a defesa dos direitos fundamentais.

Destarte, é de lucidez ímpar sobre a magnitude da influência hodierna do *mix* de dados pessoais, computação, IA, tecnologia, plataformas etc., o exemplo trazido por DIEGO MACHADO (2023), onde se pode perceber a

vulnerabilidade de direitos fundamentais dos utilizadores em questões comezinhas da vida cotidiana.

Segundo o exemplo suprarreferido, imaginemos uma pessoa que faz a locação de um veículo, onde a locadora durante a fase pré-contratual já tem acesso aos dados necessários para saber se o consumidor atende aos requisitos necessários para determinado aluguel, ofertando apenas determinadas opções de automóveis e condições de pagamentos de acordo com o histórico financeiro; que por sua vez fora inferida a partir de um modelo estatístico, usado em sistema algorítmicos de *bureau* de crédito, aos dados pessoais financeiros, de histórico de adimplemento e de crédito coletados, inclusive de terceiros com quem possui parentesco; formado o contrato, sua execução se dá através de sistemas de IoT (Internet das Coisas), com monitoramento contínuo pela empresa de locação, através de geolocalização e com sistemas de desligamento remoto para casos de inadimplência; além disso, a plataforma provedora do sistema de IoT se conecta com a plataforma de *smartphone* do usuário, permitindo uma interoperabilidade, cruzando dados pessoais e refinando o perfil, sugerindo rotas, dando comandos para produtos conectados em sistemas de *smarthome*, como aquecimento de comidas, preparação de jantar, climatização do ambiente, tudo calculado para que o usuário possa chegar em casa e ter tudo ao alcance, sem desperdício de tempo (Machado, 2023, p. 32).

Essa realidade já está presente em nosso dia a dia, e todo esse sistema, basicamente, move-se com dados (pessoais) e uso de sistemas de IA num ambiente lacunoso quanto ao *enforcement* de direitos fundamentais, colocando em voga questões constitucionais e éticas.

São muitos os escândalos recentes sobre a desregulada utilização de dados, algoritmos, IA e plataformas digitais, como podemos constatar através de casos notórios, como por exemplo a manipulação de votos em eleições *(Cambridge Analytica)*[26]*;* ferramentas de recrutamento que discriminam minorias (*Amazon Recruiting Tools)*[27]*; software* concebido para identificar a possibilidade

[26] Para maiores informações, acessar: https://www.bbc.com/portuguese/internacional-4346175;
[27] Para maiores informações, acessar: https://www1.folha.uol.com.br/tec/2018/10/amazon-desiste-de-ferramenta-de-recrutamento-que-penalizava-mulheres.shtml;

de reincidência futura de detento *(Compass*)[28]; ou para modificar o preço de acordo com a procedência do cliente (*geopricing*)[29].

À evidência, o termo algoritmo pode surpreender, vez que não tem sua gênese ligada a questões tecnológicas e de informática, pelo contrário, remonta à ideia de designação de regra de ação clara, utilizada para resolver problemas em etapas individuais definidas, ou seja, é a utilização da lógica para listar passos ordenados que resultam na solução de um determinado problema.

Algoritmos, portanto, são indispensáveis em quase todas as áreas da sociedade, mas, especialmente, para a comunicação digital e o funcionamento das modernas infraestruturas de comunicação, como a internet. Para que sejam utilizados em computadores, os algoritmos são "desenhados" em linguagem digital processável por máquina e a respectiva tarefa é processada com a ajuda de um número de etapas individuais predefinidas[30].

Todavia, da má utilização dos algoritmos poderão resultar variados problemas, nomeadamente éticos e jurídicos, como a discriminação algorítmica e a manipulação comportamental, tendo em vista que algoritmos conseguem mudar a nossa percepção do mundo e afetar comportamentos, influenciando decisões a partir de dados.

Nesta senda, podemos dizer que a Inteligência Artificial (IA) é um campo de pesquisa dentro ciência da computação e ciência da computação em geral. O termo Inteligência Artificial foi cunhado por Stanford Professor Honorário da Universidade JOHN MCCARTHY (2007), definindo-a como:

> "The science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable." (McCarthy, 2007, p.2)[31]

[28]Para maiores informações, acessar: https://www.abc.net.au/radionational/programs/lawreport/algorithms-in-the-justice-system/7676710;

[29] Para maiores informações, acessar:https://oglobo.globo.com/boa-viagem/entenda-que-geopricing-como-hoteis-no-exterior-podem-estar-cobrando-mais-caro-de-brasileiros-1-25077743;

[30] Idem. p. 12.

[31] É a ciência da engenharia de fabricar máquinas inteligentes, especialmente programas de computador inteligentes. Está relacionado à tarefa semelhante de usar computadores para

Igualmente, muito relevante é a conceituação trazida pelo Grupo de Peritos de Alto Nível sobre Inteligência Artificial no documento “Uma definição de IA: Principais capacidades e disciplinas científicas”:

> [...] os sistemas de inteligência artificial (IA) são sistemas de software (e eventualmente também de hardware) concebidos por seres humanos, que, tendo recebido um objetivo complexo, atuam na dimensão física ou digital percepcionando o seu ambiente mediante a aquisição de dados, interpretando os dados estruturados ou não estruturados recolhidos, raciocinando sobre o conhecimento ou processando as informações resultantes desses dados e decidindo as melhores ações a adotar para atingir o objetivo estabelecido. Os sistemas de IA podem utilizar regras simbólicas ou aprender um modelo numérico, bem como adaptar o seu comportamento mediante uma análise do modo como o ambiente foi afetado pelas suas ações anteriores. Enquanto disciplina científica, a IA inclui diversas abordagens e técnicas, tais como a aprendizagem automática (de que a aprendizagem profunda e a aprendizagem por reforço são exemplos específicos), o raciocínio automático (que inclui o planeamento, a programação, a representação do conhecimento e o raciocínio, a pesquisa e a otimização) e a robótica (que inclui o controlo, a perceção, os sensores e atuadores, bem como a integração de todas as outras técnicas em sistemas ciberfísicos) (Grupo Independente de Peritos de alto nível sobre inteligência artificial, 2018). [32]

Outrossim, não podemos falar em IA sem recordar do professor e pesquisador ALAN TURING. O renomado cientista é uma figura imponente em ciência da computação e IA e frequentemente chamado o “pai da IA”, vez que em 1936, ele escreveu um artigo chamado “On Computable Numbers”, onde neste *paper* foram apresentados os conceitos básicos de um computador, que ficou conhecido como *Turing Machine* e foi determinante para evolução da ciência dos algoritmos e, por consequência, da IA.

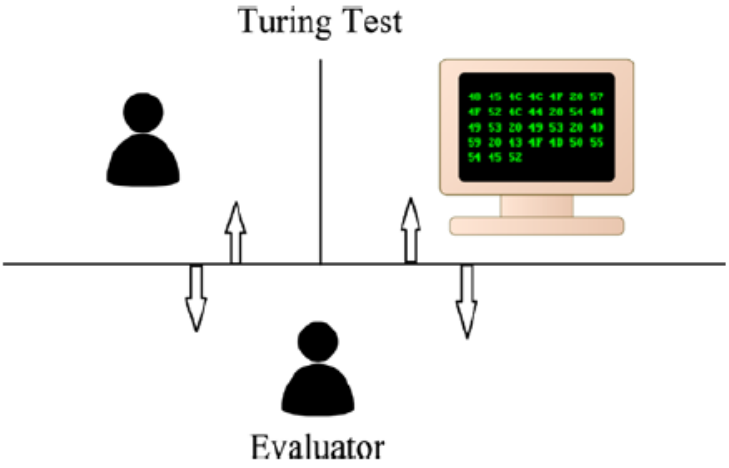
No entanto, foi através do trabalho chamado “Computing Machinery and Intelligence”, que TURING tornara-se emblemático para a IA, pois se concentrou no conceito de uma máquina que era inteligente e, para fazer isso

compreender a inteligência humana, mas a IA não precisa se limitar a métodos que sejam biologicamente observáveis. (tradução livre)

[32] Conf. em https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai; acessado em 04 de dezembro de 2023.

possível, tinha que haver uma maneira de medir e gerir, nascendo o igualmente famoso “Teste de Turing”.

O referido teste, essencialmente, traz um jogo com três jogadores: dois humanos e um computador. O avaliador, um humano, faz perguntas abertas aos outros dois (um humano e um computador) com o objetivo de determinar qual deles é o humano. Se o avaliador não pode fazer uma determinação, então presume-se que o computador é inteligente. A genialidade desse conceito é que não há necessidade de ver se a máquina realmente sabe algo, é autoconsciente ou mesmo se está correto. Em vez disso, o Teste de Turing indica que uma máquina pode processar grandes quantidades de informação, interpretar a fala e comunicar com os humanos (Taulli, 2019).



Dando um salto histórico- haja vista o recorte metodológico desta pesquisa-, a ciência da computação só fez evoluir até hoje, compreendendo, aglutinando e refinando estes fenômenos, de tal sorte que fazem parte do espectro científico de IA, atualmente, os ramos de: i) Aprendizagem Automática (*Machine Learning*); ii) Aprendizagem Profunda (*Deep Learning*); Análise Preditiva (*Predictive Analytics*); e iv) processamento de Linguagem Natural (*Natural Language Processing*).

A Aprendizagem Automática (*Machine Learning ou ML)* é do ato de usar algoritmos, técnicas, dados e procedimentos, segundo exemplos predefinidos e não mais a partir da lógica “se A, então B” (*inputs-outputs)*, habilitando o sistema de IA para resolução de problemas, sem que a máquina seja expressamente programada. Já a Aprendizagem Profunda (*Deep Learning*

ou DL) está contida no *Machine Learning (ML)* e significa aprendizagem profunda com base em redes neurais artificiais que se adaptam e aprendem a partir de um enorme conjunto de dados (*big data*[33])[34].

Destarte, relata TOM TAULLI que, após passagens pelo Massachusetts Institute of Technology (MIT) e *Bell Telephone Laboratories*, ARTHUR L. SAMUEL ingressou na IBM em 1949 no Laboratório Poughkeepsie. Seus esforços ajudaram a impulsionar o poder de computação das máquinas da empresa, como com o desenvolvimento do "701" (este foi o primeiro sistema de computador comercializado pela IBM). Mas ele também programou aplicativos - e havia um que faria história - isto é, seu jogo de damas no computador-, pois foi o primeiro exemplo de sistema de aprendizado de máquina (Taulli, 2019).

Ao olhar para o jogo de damas, TOM TAULLI mostrou como funciona o aprendizado de máquina – em outras palavras, um computador poderia aprender e melhorar processando dados sem ter que ser explicitamente programado. Isto foi possível aproveitando conceitos avançados de estatística, especialmente com análise de probabilidade. Assim, um computador poderia ser treinado para fazer previsões precisas (Taulli, 2019).

A Análise Preditiva (*Predictive Analytics),* desta forma, é o processo de usar dados para prever resultados futuros. O processo usa análise de dados, *Machine Learning*, inteligência artificial e modelos estatísticos para encontrar padrões que possam prever comportamentos futuros. As organizações podem usar dados históricos e atuais para prever tendências e comportamentos com segundos, dias ou anos de antecedência, com muita precisão[35].

Nesse sentido para Hildebrandt *apud* Machado (2023) a *computação preemptiva*, entendida como um tipo de computação que combina análise preditiva com intervenções computacionais destinadas a substituir a ação

---

[33] Em um relatório da *International Data Corporation* (IDC) chamado "Data Age 2025", a quantidade de dados criados espera-se que atinja impressionantes 163 zetabytes até 2025. Isso é cerca de dez vezes o valor em 2017. Disponível em blog.seagate.com/business/enormous-growth-in-data-is-coming-how-to-prepare-for-it-and-prosper-from-it/. Acesso em 04 de dezembro de 2023

[34]Conferir IA e machine learning: quais as diferenças?. Disponível em cloud.google.com/learn/artificial-intelligence-vs-machine-learning?hl=pt-br. Acesso em 04 de dezembro de 2023.

[35] Disponível em https://cloud.google.com/learn/what-is-predictive-analytics?hl=pt-br; Acesso em 05 de dezembro de 2023.

humana, atendendo-as ou anulando-as antes que o humano tenha chance de formar uma intenção consciente, possibilita de forma exponencial a normatividade tecnológica ou *rule of tech*, o que me muitas situações afronta o próprio Estado de Direito.

Ao seu turno, o Processamento de Linguagem Natural (*Natural Language Processing* ou PLN*)* refere-se a uma área dentro da inteligência artificial que auxilia os computadores a entender, interpretar e reproduzir a linguagem humana, permitindo, por exemplo, que a tecnologia entenda uma pessoa falando com ela. Além de entender essa linguagem, o PLN também capacita os dispositivos para criar respostas, seja por meio de textos ou áudios. O PLN compreende, interpreta e simula a linguagem natural das pessoas, promovendo uma conversação e interação bastante semelhantes à que acontece entre dois seres humano[36]. Exemplos: SIRI (IoS) e Alexa (Amazon)

Outrossim, no que se refere aos tipos de Inteligência Artificial, fala-se em *i) Inteligência Artificial Débil; Fraca(weak) ;Estreita*, para caracterizar um tipo de IA centrada em uma determinada tarefa, limitada a uma aplicação específica, ou seja, reproduz uma tarefa restrita par qual foi concebida (Ex: *Deep Blue* da IBM foi uma máquina cuja função única era jogar xadrez); e *ii) Inteligência Artificial Forte/Profunda***,** criada para reproduzir a complexidade do pensamento humano, com consciência ou autoconsciência, sendo definida como a capacidade de raciocinar, representar o conhecimento, planejar, aprender, comunicar-se me qualquer linguagem natural, e integrar todas essas habilidades para um objetivo comum.[37]

2.2) Riscos e modelos de IA :

A Inteligência Artificial talvez seja o assunto mais em voga no momento da sociedade atual, ao lado das mudanças climáticas (Hacker, 2023a). Alguns autores, como GEOVANNI DE GREGORIO, passam a denominar este

[36] Disponível em https://www.inbot.com.br/chatbots/pln-processamento-de-linguagem-natural/o-que-e-pln/?utm_source=search&utm_medium=cpc&utm_campaign=google_ads_inbot_blog&utm_id=gads_blog_inbot&gclid=CjwKCAjw_aemBhBLEiwAT98FMnSFzv6c_ti9VkHnyeDpi_ZlxDI2s88ZeFAVAizlQ0MGNAkr5ydPshoCUlgQAvD_BwE; Acessado em 02 de agosto de 2023

[37] Prevalece o entendimento que, até hoje, não houve um sistema capaz de reproduzir a consciência humana.

fenômeno da vida fulcrada e determinada por sistemas de IA não mais como o Estado de Direito (*rule of law*), mas como "Estado sob domínio Tecnológico" (*rule of tech*) ou "Estado sob regras das plataformas (*rule of platforms*), em razão das interferências que estes sistemas implicam na vida humana, produzindo normatividade como se fossem legitimadas para a função – o que não se vislumbra, à margem das balizas constitucionais e de tratados de direitos humanos.

No escólio da Professora RAQUEL BRÍZIDA a referida normatividade digital e as relações entre os sistemas tecnológico e digital e o sistema jurídico configuram, atualmente, verdadeiros paradoxos para o Direito Público, em especial, para o Direito Constitucional, na medida em que:

> [...]Nos últimos anos, assistimos a uma implacável desconstrução da ordem jurídico-constitucional, perpetrada pelos factos carreados para o sistema normativo pela Internet e Novas Tecnologias. [...] complicados "puzzle", fundado num inelutável Pluralismo Normativo Multinível – Plurinormativismo Tecnológico e Digital. Como vimos, dissonante dos pilares tradicionais do Direito Constitucional, mas particularmente atrativo no ciberespaço. (Castro, 2023, p.109)

Nesta toada, conclui-se que o papel do constitucionalismo digital ante a "constelação de novas perplexidades constitucionais", é de recuperar margem concreta de atuação na defesa do princípio da proteção dos direitos fundamentais *by default e by design*, ou seja, desde a concepção, durante todo a trajetória de tratamento e por defeito, através de critérios de transparência e de justiça algorítmica, ante aos riscos proeminentes advindos da utilização de IA (Castro, 2023).

Tanto o RGPD[38], quanto a LGPD[39] brasileira, tem como aspecto fundamental o princípio da "privacy by design and default" que importa proteger, desde a concepção de um produto ou serviço, mantida em todo o ciclo de vida do mesmo, entre controladores e operadores e titulares dos dados pessoais, como na obtenção do consentimento, na determinação de mecanismos de controle de dados e na delimitação das finalidades e do escopo do

[38] Conf. art. 25.º do RGPD.
[39] Conf. art. 46,§2º da LGPD.

processamento, de maneira que o planejamento atento à privacidade preceda o processamento de dados pessoais (Arbix, 2020, p. 48).

Em termos de riscos, para além dos jurídicos – objeto de análise mais detida nesta investigação-, não podemos deixar de tecer breves comentários sobre os riscos ligados também ao campo da ética. Todavia, ressalva-se, desde já, que não será objeto de maiores aprofundamentos, dada a limitação dessa dissertação, ainda que reconheçamos o papel fundamental do campo da ética dentro da abordagem holística de regulação de IA, advogada pela corrente doutrinária do constitucionalismo digital.

Podemos listar como problemas de natureza ética enfrentados pela sociedade algorítmica com o uso desregulado de IA àqueles relacionados aos vieses algorítmicos, à privacidade e proteção dos dados pessoais, à transparência e explicabilidade dos sistemas de IA, vigilância e controle, impactos no mercado de trabalho, dentre outros.

A ética, pois, é uma disciplina acadêmica que constitui um subdomínio da filosofia. Em termos gerais, trata de questões como “O que é uma boa ação?”, “Qual é o valor de uma vida humana?”, “O que é a justiça?” ou “O que é uma boa vida?”. Na ética académica, há quatro grandes domínios de investigação:

> i) metaética, que diz sobretudo respeito ao significado e à referência das frases normativas, e à questão de saber como os seus valores de verdade podem ser determinados (caso existam), ii) ética normativa, um meio prático para determinar a orientação moral a seguir, mediante um exame das normas de boa e má ação e da atribuição de um valor a ações específicas, iii) ética descritiva, que visa fazer uma investigação empírica do comportamento e das convicções morais das pessoas, iv) ética aplicada, respeitante ao que somos obrigados (ou autorizados) a fazer numa situação específica (muitas vezes historicamente nova) ou num determinado domínio (muitas vezes sem precedentes históricos) de possibilidades de ação (Castro, 2023, p. 113)

A ética aplicada trata de situações da vida real, em que as decisões têm de ser tomadas sob pressão do tempo e muitas vezes com uma racionalidade limitada. A ética da IA é geralmente encarada como um exemplo de ética aplicada e centra-se nas questões normativas suscitadas pela

concepção, pelo desenvolvimento, pela implantação e pela utilização da inteligência artificial[40].

Na esteira do defendido por SATOR *apud* CASTRO (2023), os direitos humanos fornecem, precisamente, um fator de conexão importante entre lei e moral, porquanto afirmam-se como exigências éticas primárias e como oportunidades, que incluem liberdades negativas e liberdades positivas. Os direitos fundamentais são, pois, direitos de caráter moral e jurídico.

Assim, uma IA de confiança deve ser i) legítima, por ser conformar aos limites constitucionais regulatórios; iii) robusta, tanto em nível técnico, quanto em nível de ambiente social e, sobretudo, iii) ética, por respeitar princípios e valores de ordem ética:

> (1) lawful – respecting all applicable laws and regulations; (2) ethical – respecting ethical principles and values; (3) robust – both from a technical perspective while taking into account its social environment[41].

Nesta senda, para além de direitos suscetíveis de proteção judicial, os direitos fundamentais devem, igualmente, serem entendidos como direitos universais, alicerçados no estatuto moral inerente aos seres humanos, e, nesse conspecto, estão também sujeitos à Ética da IA, respeitante a normas éticas que, embora não sejam necessariamente vinculativas em termos jurídicos, são cruciais para assegurar a fiabilidade do sistema (Castro, 2023).

Logo, existem quatro princípios éticos que devem ser respeitados pelos sistemas de IA baseado em confiança -imperativos éticos que os profissionais no domínio da IA devem esforçar-se sempre por respeitar-: i)

[40] Conf. Ethics guidelines for trustworthy AI – High-Level Expert Group on AI presented Ethics Guidelines for Trustworthy Artificial Intelligence, disponível em https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai, Acessado em 04 de dezembro de 2023.
[41] Idem.

respeito da autonomia humana[42]; ii) prevenção de danos[43]; iii) equidade[44]; iv) explicabilidade[45].

De mais, os princípios acima descritos devem se somar à sete requisitos essenciais em ordem de concretizar um sistema de IA de confiança, quais sejam, i) ação e supervisão humana[46]; ii) solidez técnica e confiança[47]; iii)

[42] Os direitos fundamentais em que a UE se alicerça visam garantir o respeito da liberdade e da autonomia dos seres humanos. Os seres humanos que interajam com sistemas de IA devem poder manter uma autodeterminação plena e efetiva sobre si próprios e participar no processo democrático. Os sistemas de IA não devem subordinar, coagir, enganar, manipular, condicionar ou arregimentar injustificadamente os seres humanos. Em vez disso, os sistemas de IA devem ser concebidos para aumentar, complementar e capacitar as competências cognitivas, sociais e culturais dos seres humanas. A distribuição de funções entre os seres humanos e os sistemas de IA devem seguir princípios de conceção centrados no ser humano e deixar uma oportunidade significativa para a escolha humana. Isto implica que se garanta a supervisão28 e o controlo por parte de seres humanos sobre os processos de trabalho dos sistemas de IA. Estes sistemas também podem alterar radicalmente a esfera do trabalho, que deverá apoiar os seres humanos no ambiente de trabalho e visar a criação de um trabalho significativo

[43] Os sistemas de IA não devem causar danos ou agravá-los nem afetar negativamente os seres humanos de qualquer outra forma. Isto implica a proteção da dignidade, bem como da integridade mental e física, do ser humano. Os sistemas de IA e os ambientes em que operam devem ser seguros e protegidos. Devem ser tecnicamente sólidos e deve garantir-se que não estão abertos a utilizações malévolas. As pessoas vulneráveis devem receber maior atenção e ser incluídas no desenvolvimento e na implantação dos sistemas de IA. Há também que prestar especial atenção às situações em que os sistemas de IA podem causar ou agravar impactos negativos devido a assimetrias de poder ou de informação, nomeadamente entre empregadores e trabalhadores, empresas e consumidores ou governos e cidadãos. A prevenção dos danos implica também ter em consideração o ambiente natural e todos os seres vivos.

[44] O desenvolvimento, a implantação e a utilização dos sistemas de IA devem ser equitativos. Embora reconheçamos que há muitas interpretações diferentes de equidade, consideramos que esta tem uma dimensão substantiva e processual. A dimensão substantiva implica um compromisso com: a garantia de uma distribuição equitativa e justa dos benefícios e dos custos, bem como de inexistência de enviesamentos injustos, discriminação e estigmatização contra pessoas e grupos. Se for possível evitar os enviesamentos, os sistemas de IA podem até aumentar a equidade societal. A igualdade de oportunidades em termos de acesso à educação, aos bens e serviços e à tecnologia deve ser igualmente promovida. Além disso, a utilização de sistemas de IA nunca deverá levar a que os utilizadores (finais) sejam iludidos ou prejudicados na sua liberdade de escolha. Além disso, a equidade implica que os profissionais no domínio da IA devem respeitar o princípio da proporcionalidade entre os meios e os fins, e analisar cuidadosamente a forma de equilibrar os interesses e objetivos em causa. A dimensão processual da equidade implica uma possibilidade de contestar e procurar vias de recurso eficazes contra as decisões tomadas por sistemas de IA e pelos seres humanos que os utilizam32. Para o efeito, a entidade responsável pela decisão deve ser identificável e os processos decisórios explicáveis.

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

[46] Ação e supervisão humanas, incluindo os direitos fundamentais, a ação humana e a supervisão humana

[47] Solidez técnica e segurança, incluindo a resiliência perante ataques e a segurança, os planos de recurso e a segurança geral, a exatidão, a fiabilidade e a reprodutibilidade

privacidade e governação de dados[48]; iv) transparência[49]; v) diversidade, não-discriminação e equidade[50]; vi) bem-estar societal e ambiental[51]; vi) responsabilidade[52].

Percebe-se que o imperativo ético elencado sobre o devido respeito à autonomia humana e ao requisito essencial da privacidade e governação de dados formam um núcleo rígido, no que tange à proteção de dados e privacidade em sistemas de IA. Da autonomia, como uma decorrência da dignidade humana, como já demonstrado anteriormente, exsurge autodeterminação, tanto informacional, quanto pessoal, que levam ao direito à proteção de dados e à privacidade, respectivamente.

É inegável que a IA, como uma família de tecnologias em rápida evolução, é capaz de oferecer um vasto conjunto de benefícios econômicos e sociais a todo o leque de indústrias e atividades sociais, notadamente ao melhorar as previsões, otimizar as operações e a afetação de recursos e personalizar o fornecimento dos serviços, contribuindo para resultados benéficos para a sociedade e o ambiente e conceder vantagens competitivas às empresas e à economia, como nos domínios das alterações climáticas, *life sciences*, finanças, políticas públicas, mobilidade e outros.[53]

Todavia, existem graves riscos que devem ser endereçados com o avanço desta tecnologia, que até então são desconhecidos ou pouco conhecidos da sociedade e academia, que merecem total atenção.

O *recitals* 1.1 do AIA toma em conta esta preocupação, expressamente salientando que os mesmos elementos e técnicas que produzem os benefícios socioeconômicos da IA também podem trazer novos riscos ou consequências negativas para os cidadãos e a sociedade e, para tanto, uma

[48] Privacidade e governação dos dados, incluindo o respeito da privacidade, a qualidade e a integridade dos dados e o acesso aos dados.
[49] Transparência, incluindo a rastreabilidade, a explicabilidade e a comunicação.
[50] Diversidade, não discriminação e equidade, incluindo a prevenção de enviesamentos injustos, a acessibilidade e a conceção universal e a participação das partes interessadas
[51] Bem-estar societal e ambiental, incluindo a sustentabilidade e o respeito do ambiente, o impacto social, a sociedade e a democracia.
[52] Responsabilização, incluindo a auditabilidade, a minimização e a comunicação dos impactos negativos, as soluções de compromisso e as vias de recurso.
[53] Nesse sentido conferir *Recitals* 1.1 do Regulamento do Parlamento Europeu e do Conselho que Estabelece Regras Harmonizadas em Matéria de Inteligência Artificial (Regulamento Inteligência Artificial) E Altera Determinados Atos Legislativos Da União – AI ACT.

regulação eficaz deverá ter em mente os valores, princípios e direitos fundamentais, pois a inteligência artificial deve ser uma ferramenta ao serviço das pessoas e uma força positiva para a sociedade com o objetivo final de aumentar o bem-estar dos seres humanos.:

> "À luz da velocidade da evolução tecnológica e dos possíveis desafios, a UE está empenhada em alcançar uma abordagem equilibrada. É do interesse da União preservar a liderança tecnológica da UE e assegurar que novas tecnologias, desenvolvidas e exploradas respeitando os valores, os direitos fundamentais e os princípios da União, estejam ao serviço dos cidadãos europeus"[54]

À evidencia, existe uma preocupação do impacto de sistemas de IA nos direitos fundamentais, especialmente pelo mal uso da inteligência artificial, como os riscos ligados à imprevisibilidade, autonomia técnica, opacidade, falta de dados técnicos para treinamento e manipulação (Hacker, 2023a). Por consequência, acabam atingindo mais fortemente a proteção de dados pessoais, a privacidade e a igualdade, no viés da discriminação levada à efeito pelos algoritmos.

Por isso, HACKER (2023a) relembra a árdua a tarefa regulatória da UE neste assunto, para tentar conseguir conceber um projeto de regulamentação que ponha em compatibilidade, constitucionalmente justificável, um *ecosystem of excellence* com um *ecosystem of trust.*

Neste sentido, o *Conseil de L'Europe,* através de seu Comitê para Inteligência Artificial, elaborou o *draft framework* para Convenção de inteligência artificial, direitos humanos, democracia e *rule of law*, onde determina claramente como obrigações gerais assumidas por sistemas de IA a i) proteção dos direitos humanos; ii) defesa da integridade do processo democrático; e iii) o respeito ao Estado de Direito[55].

De mais a mais, elenca expressamente os princípios nos quais os sistemas de IA deverão estar ancorados, sendo eles: i) Dignidade da Pessoa

[54] Idem.
[55] Conf. Draft framework convention on artificial intelligence, human rights, democracy and the rule of law. Comitê para Inteligência Artificial, 2023. p.7

Humana e sua autonomia; ii) transparência e supervisão; iii) accountability e responsabilização; iv) igualdade e não-discriminação; v) privacidade e proteção de dados pessoais; vi) sustentabilidade ambiental e preservação da saúde; vii) precaução e confiança; e viii) inovação segura[56].

Nesse sentido, o Regulamento para Inteligência Artificial (AIA) estabelece regras harmonizadas para o desenvolvimento, a colocação no mercado e a utilização de sistemas de IA na União, na sequência de uma abordagem proporcionada baseada no risco (*risk-based*). O AIA, servindo de inspiração para outras legislações - como a brasileira-, optou uma abordagem baseada em risco que faz diferenciações entre as utilizações de IA que possam criam: i) um risco inaceitável, ii) um risco elevado, iii) um risco baixo ou mínimo, com condutas já proibidas *a priori*, consoante se lê do artigo n.5º, como por exemplo:

> 1.a) A colocação no mercado, a colocação em serviço ou a utilização de um sistema de IA que empregue técnicas subliminares que contornem a consciência de uma pessoa para distorcer substancialmente o seu comportamento de uma forma que cause ou seja suscetível de causar danos físicos ou psicológicos a essa ou a outra pessoa.

Dessarte, na lista de práticas proibidas inclui-se todos os sistemas de IA cuja utilização seja considerada inaceitável por violar os valores, princípios e direitos fundamentais. As proibições abrangem práticas com potencial significativo para manipular as pessoas por meio de técnicas subliminares que lhes passam despercebidas ou explorar as vulnerabilidades de grupos específicos, como as crianças ou as pessoas com deficiência, para distorcer substancialmente o seu comportamento de uma forma que seja suscetível de causar danos psicológicos ou físicos a essa ou a outra pessoa[57].

Outras práticas manipuladoras ou exploratórias que são possibilitadas pelos sistemas de IA e que afetam os adultos podem ser abrangidas pela legislação em matéria de proteção de dados, de defesa dos consumidores e de serviços digitais, que garante que as pessoas singulares sejam devidamente

[56] Ibidem.
[57] Conf. *Recitals* 5.2.2 do AI ACT.

informadas e tenham a liberdade de decidir não se sujeitar a uma definição de perfis ou a outras práticas que possam afetar o seu comportamento[58].

A proposta também proíbe sistemas de classificação social (*social scoring systems*), assente na IA, para uso geral por parte das autoridades públicas. Por último, é igualmente proibida a utilização de sistemas de identificação biométrica à distância em tempo real (*long-distance biometric ID*), em espaços acessíveis ao público para efeitos de manutenção da ordem pública, a não ser que se apliquem determinadas exceções limitadas.[59]

Por sua vez, os sistemas de IA de risco elevado estão afetos a regras específicas em domínios como a saúde e a segurança ou para os direitos fundamentais de pessoas singulares. Logo, em conformidade com uma abordagem baseada no risco, esses sistemas de IA de risco elevado são autorizados em determinado local, mas estão sujeitos ao cumprimento de determinados requisitos obrigatórios e a uma avaliação da conformidade *ex ante.*

Ou seja, existem requisitos específicos para sistemas de IA de risco elevado e obrigações para os operadores desses sistemas, considerando como risco elevado quando o sistema de IA destina-se a ser utilizado como um componente de segurança de um produto ou é, ele próprio, um produto abrangido pela legislação de harmonização da União enumerada no anexo II[60].

Para esses sistemas de risco elevando deve-se cumprir obrigações mais severas, descritas no “Capitulo 2” da proposta, tais como i) criação e implementação de gestão de riscos; ii) governação de dados; ii) documentação técnica; iv) manutenção de registos; v) transparência; vi) prestação de informações aos utilizadores; vii) cibersegurança; e viii) supervisão humana[61].

---

[58] Idem.

[59] Ibidem.

[60] Nos termos da legislação de harmonização da União enumerada no anexo II, o produto cujo componente de segurança é o sistema de IA, ou o próprio sistema de IA enquanto produto deve ser sujeito a uma avaliação da conformidade por terceiros com vista à colocação no mercado ou à colocação em serviço

[61] Quanto à supervisão humana, em linha com que já indicamos no início deste capítulo relativo a sua fundamentalidade num Estado de Direito, como decorrência da própria dignidade humana, trata-se de uma regra deveras importante, vez que impõe aos sistemas de IA de risco elevado que devam ser concebidos e desenvolvidos de tal modo, incluindo com ferramentas de interface homem-máquina apropriadas, que possam ser eficazmente supervisionados por pessoas singulares durante o período de utilização do sistema de IA, a fim de prevenir ou minimizar os riscos para a saúde, a segurança e os direitos fundamentais que possam surgir quando um sistema de IA e risco elevado é usado em conformidade com a sua finalidade prevista ou em condições de utilização indevida razoavelmente previsíveis.

Para o AIA na UE, um sistema de IA é considerado de risco elevado quando estejam satisfeitas ambas as condições que se seguem: a) o sistema de IA destina-se a ser utilizado como um componente de segurança de um produto ou é, ele próprio, um produto abrangido pela legislação de harmonização da União enumerada no anexo II; b) nos termos da legislação de harmonização da União enumerada no anexo II, o produto cujo componente de segurança é o sistema de IA, ou o próprio sistema.

São igualmente considerados de risco elevando àqueles sistemas de IA elencados no anexo III da proposta: i) identificação biométrica e categorização de pessoas singulares; ii) gestão e funcionamento de infraestruturas críticas; iii) educação e formação profissional; iv) emprego, gestão de trabalhadores e acesso ao emprego por conta própria; v) acesso a serviços privados e a serviços e prestações públicas essenciais, bem como o usufruto dos mesmos; vi) manutenção da ordem pública; vii) gestão da migração, do asilo e do controlo das fronteiras; e viii) administração da justiça e processos democráticos.

Os sistemas autônomos, que muito interessam à regulação dada sua capacidade de tomada de decisões sem a participação humana, classificam-se pela legislação da UE, como de risco elevado, ou seja, sistemas de IA de risco elevado que não são componentes de segurança de produtos nem são, eles próprios, produtos, é apropriado classificá-los como de risco elevado se, em função da finalidade prevista, representarem um risco elevado de danos para a saúde e a segurança ou de prejuízo para os direitos fundamentais das pessoas, tendo em conta a gravidade dos possíveis danos e a probabilidade dessa ocorrência, e se forem utilizados num conjunto de domínios especificamente predefinidos no regulamento.

Em nítida inspiração do modelo europeu, no âmbito da regulamentação da matéria no Brasil, o Projeto de Lei (PL) n.º2338, de 2023 optou também por uma abordagem regulatória baseada e riscos (*risk-based approach*). Estes são subdivididos em i) Riscos Excessivos (Seção II) e ii) Alto Risco (Seção III), encaixando-se os sistemas autônomos como de natureza de alto risco.

Outrossim, são elencados domínios onde expressamente não se poderá utilizar sistemas de IA, tais como a) que empreguem técnicas subliminares que tenham por objetivo ou por efeito induzir a pessoa natural a se comportar de forma prejudicial ou perigosa à sua saúde ou segurança ou contra os fundamentos desta Lei; e b) que explorem quaisquer vulnerabilidades de grupos específicos de pessoas naturais, tais como as associadas a sua idade ou deficiência física ou mental, de modo a induzi-las a se comportar de forma prejudicial a sua saúde ou segurança ou contra os fundamentos desta Lei.

A Inteligência Artificial Generativa, por sua vez, são sistemas de IA que podem criar novos conteúdos, sendo que os modelos mais populares geram texto e imagens a partir de *prompts* de texto, mas alguns usam outras entradas, como imagens, para criar áudio, vídeo e imagens[62]. Historicamente, o *machine learning* era amplamente limitado a modelos preditivos, usados para observar e classificar padrões em conteúdo, o que se denomina de IA não-generativa. A IA generativa, portanto, foi um avanço por não se limitar a lógica de entender e classificar o padrão; o aprendizado de máquina de natureza generativa é capaz de criar novos padrões sem *inputs* humanos.

Portanto, com a compreensão adquirida, a IA generativa se torna capazes de gerar e criar novos dados, semelhantes aos utilizados na formação inicial do padrão, sem a necessária intervenção prévia humana na classificação e padronização dos dados. O avanço deste tipo de modelo levou a tecnologia para a fronteira do conhecimento, dando um largo passo para uma possível consciência ou um modelo de IA Forte.

A IA generativa usa um modelo de ML para aprender os padrões e as relações em um conjunto de dados de conteúdo criado por humanos. Em seguida, ele usa os padrões aprendidos para gerar novo conteúdo. A maneira mais comum de treinar um modelo generativo de IA é usar o aprendizado supervisionado: o modelo recebe um conjunto de conteúdo criado por humanos

[62] Conf. HM Government *Safety and Security Risks of Generative Artificial Intelligence to 2025,* disponível em https://assets.publishing.service.gov.uk/media/653932db80884d0013f71b15/generative-ai-safety-security-risks-2025-annex-b.pdf

e rótulos correspondentes. Em seguida, ele aprende a gerar conteúdo semelhante ao criado por humanos e rotulado com os mesmos rótulos.

A IA generativa, pois, representa um avanço significativo na tecnologia e, assim, oferta grandes benefícios também. Esta processa uma vasta gama de dados, criando *insights* e respostas por meio de texto, imagens e formatos amigáveis, sendo amplamente utilizada, por exemplo, para geração de imagens, arte e texto; criação de músicas originais; simulação e treinamento complexos, especialmente com criação de dados sintéticos; conversação e *chatbots*; e medicina e pesquisa científica, especialmente na indústria farmacêutica, ao possibilitar simulação de experimentos e criação de novos medicamentos.

Existem modelos mais usados de IA generativa, como as Redes Generativas Adversariais (GANs); as Redes Neurais Recorrentes (RNNs); e Redes Neurais Recorrentes (RNNs); *Transformadores*; *Variate Autoencoders* (VAEs); e Modelos de Linguagem Autorregressivos. As Redes Generativas Adversariais (GANs) consistem em dois componentes principais, quais sejam, um gerador e um discriminador. O gerador cria dados, enquanto o discriminador avalia a autenticidade desses dados; o treinamento envolve uma competição entre esses dois componentes, melhorando constantemente a capacidade do gerador de criar dados indistinguíveis dos dados reais.

As Redes Neurais Recorrentes (RNNs) são utilizadas em tarefas sequenciais, como geração de texto; elas têm uma memória interna que lhes permite lembrar informações anteriores, sendo úteis na criação de sequências de dados coerentes e contextuais.

Os modelos de *transformers*, como o GPT (*Generative Pre-trained Transformer*), são projetados para processar dados em paralelo, tornando-os eficazes para tarefas de linguagem natural. Já os VAEs são modelos generativos que utilizam uma abordagem de codificação e decodificação para aprender a representação latente de dados. Eles são frequentemente usados para gerar novas instâncias de dados a partir dessa representação latente.

Por fim, os Modelos de Linguagem autorregressivos geram sequências de dados um elemento de cada vez, condicionados aos elementos gerados anteriormente. O exemplo mais conhecido é o GPT, que gera texto de maneira autorregressiva.

À vista disso, esses modelos são aplicados em diversas áreas, como geração de texto, imagem, áudio, e até mesmo na criação de conteúdo multimídia. Eles têm sido usados para criar arte generativa, gerar músicas, produzir imagens realistas, entre outras aplicações criativas. No entanto, é importante notar que a criação de IA generativa também levanta diversos riscos.

Segundo o relatório *Safety and Security Risks of Generative Artificial Intelligence to 2025, do* governo do Reino Unido, é provável que a IA generativa amplifique os riscos existentes do que crie riscos totalmente novos, aumentando drasticamente a velocidade e a escala de algumas ameaças. A dificuldade de prever avanços tecnológicos cria um potencial significativo de surpresa tecnológica, sendo altamente provável que surgirão ameaças adicionais que não foram previstas no próprio relatório.

A rápida proliferação e a crescente acessibilidade destas tecnologias permitirão quase certamente que agentes de ameaças menos sofisticados conduzam ataques anteriormente inatingíveis, sendo riscos classificados como de natureza i) esfera digital; ii) sistemas políticos e as sociedades; iii) segurança física

Os riscos na esfera digital, como por exemplo, ataques cibernéticos, fraudes, burlas, falsificação de identidade, imagens de abuso sexual de crianças, têm maior probabilidade de se manifestar e de ter o maior impacto até 2025.

A probabilidade de os riscos para os sistemas políticos e as sociedades aumentarem à medida que a tecnologia se desenvolve e a sua adoção se alarga. Ou seja, a proliferação de meios de comunicação sintéticos corre o risco de minar o envolvimento democrático e a confiança pública nas instituições governamentais.

Os riscos de segurança física provavelmente aumentarão à medida que a IA generativa for incorporada em mais sistemas físicos, incluindo infraestruturas críticas.

Os riscos mais significativos poderão se manifestar da seguinte forma, à teor do relatório *Safety and Security Risks of Generative Artificial Intelligence to 2025*[63]:

Ataques cibernéticos: a IA generativa pode ser usada para criar invasões cibernéticas mais rápidas, eficazes e em maior escala por meio de métodos de *phishing* personalizados ou replicação de malware.

Aumento de vulnerabilidades digitais: a integração generativa de IA em funções e infraestruturas críticas apresenta uma nova superfície de ataque através da corrupção de dados de treinamento ('envenenamento de dados'), sequestro de saída do modelo ('injeção imediata'), extração de dados de treinamento sensíveis ('inversão de modelo'), classificação incorreta de informações ('perturbação ') e visando o poder da computação[64].

Erosão da confiança na informação: a IA generativa poderá levar à poluição do ecossistema de informação pública com *bots* hiper-realistas e meios de comunicação sintéticos (*deepfakes*) que influenciam o debate social e refletem preconceitos sociais pré-existentes. Este risco inclui a criação de notícias falsas, a desinformação personalizada, a manipulação dos mercados financeiros e o enfraquecimento do sistema de justiça criminal. Até 2026, os meios de comunicação sintéticos poderão abranger uma grande proporção de conteúdo online e correm o risco de minar a confiança do público no governo, ao mesmo tempo que aumentam a polarização e o extremismo[65].

Influência política e social: já foi demonstrado que as ferramentas generativas de IA são capazes de persuadir os seres humanos sobre questões políticas e podem ser utilizadas para aumentar a escala, o poder de persuasão

[63] Disponível em https://assets.publishing.service.gov.uk/media/653932db80884d0013f71b15/generative-ai-safety-security-risks-2025-annex-b.pdf.
[64] Ibidem.
[65] Ibid.

e a frequência da desinformação e da desinformação. De forma mais geral, a IA generativa pode gerar conteúdo hiperdirecionado com escala e sofisticação sem precedentes[66].

Instrução de arma: a IA generativa pode ser utilizada para reunir conhecimentos sobre ataques físicos perpetrados por intervenientes violentos não estatais, incluindo armas químicas, biológicas e radiológicas.[67]

Uso inseguro e uso indevido: a integração generativa da IA em sistemas e infraestruturas críticas corre o risco de fugas de dados, sistemas tendenciosos e discriminatórios ou comprometimento da tomada de decisões humanas devido a uma segurança da informação deficiente e a processos algorítmicos opacos, por exemplo, “alucinações”. O uso inadequado por qualquer organização de grande escala pode ter consequências indesejadas e resultar em falhas em cascata. A integração generativa da IA em funções críticas também pode resultar numa dependência excessiva de cadeias de abastecimento que são opacas, potencialmente frágeis e controladas por um pequeno número de empresas[68].

Ainda dentro da categoria de “uso inseguro e uso indevido” de sistemas de IA generativa, de particular importância são os riscos em torno das chamadas “alucinações”, que se refere a um fenômeno em que sistemas de IA, como modelos de linguagem, produzem informações incorretas, fictícias ou que não estão fundamentadas nos dados de *input*, criando literalmente conteúdo.

À evidência, esses sistemas têm condições de produzir conteúdo, mas nem sempre garantem a verdade e precisão das informações geradas. Os modelos de IA generativa, como o ChatGPT, são treinados em grandes volumes de dados textuais para aprender padrões e estruturas linguísticas, pelo que usam esses padrões para fazer previsões sobre quais palavras ou frases provavelmente virão a seguir em uma determinada sequência de texto. No entanto, devido à complexidade da linguagem e à natureza diversificada dos

[66] Ibid.
[67] Ibid.
[68] Ibid.

dados de treinamento, esses modelos podem ocasionalmente produzir saídas que não correspondem à realidade ou que extrapolam o que está presente nos dados originais.

Nesse sentido, é importante destacar o famoso caso noticiado pelo jornal The New York Times[69] acerca da "atividade inventiva" de jurisprudência pelo ChatGPT, utilizado numa demanda em face um companhia aérea, onde o advogado, sem conferir a existência dos precedentes, colacionou nos autos documentos e decisões inventada pelo sistema de IA; bem como o caso investigado pelo Conselho Nacional de Justiça do Brasil (CNJ), onde uma sentença assinada por um magistrado federal teria sido elaborada por meio do ChatGPT que, igualmente, inventou jurisprudência para sustentar a decisão[70].

Existem diversas causas para as "alucinações", como a extrapolação criativa; lacunas nos dados de treinamento; conflitos de informação; dificuldade de contextualização[71]:

Na extrapolação criativa os modelos podem gerar informações fictícias que parecem plausíveis, mas não estão baseadas em dados reais, pois os modelos tentam prever continuamente as palavras subsequentes, muitas vezes fazendo inferências e suposições com base em padrões aprendidos durante o treinamento; já nas lacunas nos dados de treinamento, se os dados de treinamento contiverem informações incompletas, imprecisas ou ambíguas, os modelos podem preencher essas lacunas com suposições incorretas; os conflitos de informação surgem aquando os dados de treinamento contêm informações contraditórias ou ambíguas e os modelos podem não ser capazes de discernir qual é a informação correta a ser usada em uma determinada situação; por fim, na dificuldade de contextualização, os modelos podem ter

[69]Conf. https://www.nytimes.com/2023/05/27/nyregion/avianca-airline-lawsuit-chatgpt.html; Acessado em 09 de janeiro de 2024.
[70] Conf. https://www.conjur.com.br/2023-nov-12/cnj-vai-investigar-juiz-que-usou-tese-inventada-pelo-chatgpt-para-escrever-decisao/. Acessado em 09 de janeiro de 2024.
[71] XAVIER, Fábio Correa. Alucinações de IA: o lado perverso da criatividade.2023. Disponível em https://www.linkedin.com/pulse/alucina%C3%A7%C3%B5es-da-ia-o-lado-perverso-criatividade-f%C3%A1bio-correa-xavier/?originalSubdomain=pt; Acesso em 09 de janeiro de 2024.

dificuldade em entender completamente o contexto de uma conversa ou tarefa, o que pode levar a respostas inadequadas ou informações incorretas.

A questão das alucinações em IA generativa são particularmente relevantes em cenários em que a precisão é crucial, como no uso pelo poder judiciário, na redação de notícias e uso médico. A capacidade dos modelos de IA de criar informações incorretas ou fictícias podem ter implicações graves e prejudicar a confiança dos usuários nesses sistemas.

Especificamente aos riscos com repercussão jurídica, aponta Hacker (2023b) que, atualmente assistimos, em tempo real, ao nascimento de uma nova geração de sistemas de IA, particularmente no domínio da IA generativa. Esses modelos oferecem enormes oportunidades e mudarão significativamente a maneira como trabalhamos, nos comunicamos. Simultaneamente, esta nova geração de sistemas de IA comporta riscos específicos que a regulamentação precisa abordar, quais sejam, a proteção de dados; a não discriminação; a qualidade de dados e resultados; moderação de conteúdo; a sustentabilidade ambiental; e a responsabilidade civil:

> In my view, the most urgent ones, in the short and medium-term, are the following six issues: data protection; non-discrimination; quality (of data and output); content moderation; environmental sustainability; and civil liability. In the longer term, we also need to prepare for a potential restructuring of the job market, with concomitant effects on tax revenue, as well as the use of AI by malicious actors (Hacker, 2023b)[72].

Não é de difícil percepção que a grande maioria dos problemas de IA, elencados ao longo deste capítulo, perpassam por dados. Ou seja, a utilização *lato sensu* de dados está no centro do debate sobre um correto, justo, legítimo e conformado uso de IA aos limites do Estado de Direito.

[72] Na minha opinião, os mais urgentes, a curto e médio prazo, são as seis questões seguintes: proteção de dados; não discriminação; qualidade (de dados e resultados); moderação de conteúdo; sustentabilidade ambiental; e responsabilidade civil. A longo prazo, também precisamos preparar-se para um potencial reestruturação do mercado de trabalho, com efeitos concomitantes sobre a fiscalidade receita, bem como o uso de IA por atores mal-intencionados. (Tradução livre). Disponível em: https://www.bundestag.de/resource/blob/949586/7158a308bd0737da83afd2d76cf27684/Stellungnahme-Hacker-ENG.pdf. Acesso em 18 de janeiro de 2024.

Dessarte, TEKI AKUETTEH afirmou na Comissão de Juristas Responsáveis por Subsidiar Elaboração do Substitutivo sobre Inteligência Artificial no Brasil que os dados lidam com outro princípio-chave, no sentido de a existência de IA e o uso dessas fazem com que direitos ligados à privacidade sejam mais urgentes por serem centrais de muitas formas:

> Essa transformação na natureza dos dados, que se dá de forma muito acelerada, faz com que certos pontos que não eram considerados tão sensíveis se tornem agora deveras sensíveis, pois a tecnologia IA pode ser utilizada, por exemplo, para acumular imagens faciais, para criar vídeos de impostores de uma pessoa, e isso tem associações, isso tem vastas implicações sociais. Há questões políticas aí também. Então, é preciso de fato desenvolver regras nessas áreas, lado a lado, de mãos dadas, garantindo que não abandonemos esses direitos base, essas necessidades dos indivíduos diante do potencial dos grandes avanços tecnológicos (Senado Federal, 2022)[73].

Portanto, endereçar os problemas advindos do uso de sistemas de IA, para que haja uma regulação robusta e constitucionalmente adequada, compreende o respeito pelos princípios estruturantes do Estado de Direitos e respeito aos direitos fundamentais, para além da observância de valores éticos, como mencionado anteriormente. Podemos comprovar tal afirmação do próprio regulamento de IA que indica como objetivos específicos garantir que os sistemas de IA colocados no mercado da União e utilizados sejam seguros e respeitem a legislação em vigor em matéria de direitos fundamentais e valores da União[74].

Em terras brasileiras, o PL n.2338/2023 tem o objetivo de proteger direitos fundamentais e garantir a implementação de sistemas seguros e confiáveis, em benefício da pessoa humana, do regime democrático e do desenvolvimento científico e tecnológico utilizando como fundamentos, dentre outros, *a)* centralidade da pessoa humana; *b)* o respeito aos direitos humanos e aos valores democráticos; *c)* livre desenvolvimento da personalidade; e *d)* a privacidade, a proteção de dados e a autodeterminação informativa.

Desse modo, é preciso encarar a proteção de dados com a devida relevância para uma regulação dos riscos apresentados pelos sistemas de IA,

[73] Conf. Relatório Final – Comissão de Juristas Responsável por Subsidiar Elaboração do Substitutivo sobre Inteligência Artificial no Brasil. Senado Federal: Brasília. 2023, p. 137.
[74] Conf. item 1.1 do Regulamento para Inteligência Artificial da UE.

também por se tratar de expressão fundamental da autonomia e autodeterminação do indivíduo e, por isso, extraída da própria dignidade da pessoa humana.

Para tanto, torna-se imperioso a observância, para além os princípios gerais de direito público os específicos para os sistemas de IA[75], àqueles princípios e fundamentos ligados propriamente à proteção de dados, como os elencados especialmente nos artigos 2.º e 6.º da Lei Geral de Proteção de Dados brasileira e no Capítulo II do Regulamento Geral de Proteção de Dados da União Europeia.

3) Proteção de dados pessoais e Inteligência Artificial (IA):

3.1) Breve histórico e evolução do direito à proteção de dados:

Historicamente, a partir do estudo comparado entre Brasil e União Europeia, percebemos que, ao longo da evolução da internet, o uso de dados pessoais e a consequente autodeterminação vem sendo reconhecido como parte central da solução dos problemas enfrentados na nova esfera pública digital. Soma-se isso a evolução da IA e todas as suas aplicações inovadoras, disruptivas, e seus desafios e riscos na vida cotidiana.

A rápida evolução tecnológica e a globalização criaram novos desafios em matéria de proteção de dados pessoais. A recolha e a partilha de dados pessoais registaram um aumento significativo e as novas tecnologias permitem às empresas privadas e às entidades públicas a utilização de dados pessoais numa escala sem precedentes no exercício das suas atividades; as pessoas disponibilizam cada vez mais as suas informações pessoais de uma forma pública e global, enquanto as novas tecnologias transformaram a economia e a vida social[76].

---

[75] i) Dignidade da Pessoa Humana e sua autonomia; ii) transparência e supervisão; iii) *accountability* e responsabilização; iv) igualdade e não-discriminação; v) privacidade e proteção de dados pessoais; vi) sustentabilidade ambiental e preservação da saúde; vii) precaução e confiança; e viii) inovação segura.

[76] Disponível em https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=celex%3A32016R0679. Acesso em 06 de julho de 2023.

Desta maneira, há uma interseção evidente entre dados pessoais e IA que, por conseguinte, desafia uma aproximação entre a regulamentação e normatização destes assuntos. Neste sentido, à teor do que consta na justificativa do PL nº 2338/2023, a proposta estabelece uma regulação baseada em riscos e uma modelagem regulatória baseada em direitos.

Parte da premissa de que não há um trade-off entre a proteção de direitos e liberdades fundamentais, da valorização do trabalho e da dignidade da pessoa humana face à ordem econômica e à criação de novas cadeias de valor. Pelo contrário, seus fundamentos e a sua base principiológica buscam tal harmonização, nos termos da Constituição Federal. Essa opção de política regulatória é acertada e traça uma importante aproximação com a normativa brasileira de proteção de dados pessoais, que também compatibiliza a abordagem baseada em direitos (*rights-based approach*) com a abordagem baseada no risco (*risk-based approach*), visto que ambas não são excludentes entre si (ANPD, 2023).

Exemplificando, o PL nº 2338/2023, a um só tempo, proíbe sistemas de IA de risco excessivo; delimita quais são os sistemas ditos de alto risco e suas obrigações e prevê direitos às pessoas naturais afetadas pelo funcionamento desses sistemas; enquanto a LGPD estabelece diferenças de regime jurídico ou de carga regulatória com base no nível de riscos gerado pela atividade de tratamento de dados executada pelo agente de tratamento. Além disso, a normativa garante um conjunto de direitos aos titulares, ainda que a atividade de tratamento implique baixo risco a direitos, garantias e liberdades fundamentais da pessoa (ANPD, 2023).

Entretanto, antes de adentrar às interseções entre proteção de dados e inteligência artificial, especialmente em relação às normativas, é importante fazer uma breve digressão histórica.

Desde o julgamento pela Corte Constitucional alemã (BvG) da Lei do Censo, que expressamente reconheceu a autodeterminação informacional, passando pelo reconhecimento como Direito Fundamental na Carta da União Europeia, até a entrada em vigor da RGPD, atestamos a importância que o espaço europeu confere ao tema da proteção de dados.

No Brasil, com certo atraso, podemos com certa segurança dizer que foi, nomeadamente, com a edição da LGPD, em 2018, e através do julgamento da ADI 6389 e conexas, em 2020, que a autodeterminação informacional passou a ser reconhecida, o que desencadeou a entrada em vigor antecipada da LGPD e a aprovação da EC 115/22, que inseriu a proteção de dados pessoais no rol do artigo 5º da Carta Maior, integrando formalmente o catálogo dos Direitos Fundamentais da República Federativa do Brasil.

À evidência, e inarredável reconhecer a proteção de dados como corolário da autodeterminação informacional, decorrência imediata da dignidade da pessoa humana.

Dessarte, vamos nos valer de dois exemplos emblemáticos que corroboraram a ideia de autodeterminação no ciberespaço, oriundos do reconhecimento pelo Poder Judiciário, como os casos brasileiros das ADI 6388 e conexas e a consequente aprovação da EC n.º115/22 e o RE n.º1.010.606 (“caso Aida Curi”) comparando com o célebre caso da Corte Constitucional alemã e a lei do Censo e o caso Google Spain x Costeja Gonzales.

Em meio a pandemia de COVID-19, o governo federal editou a Medida Provisória nº 954/2020, que dispunha sobre o compartilhamento obrigatório, pelas empresas de telecomunicações, de dados pessoais dos usuários com o Instituto Brasileiro de Geografia e Estatística (IBGE), a fim de produzir as *estatísticas oficiais do país*, a par de uma Lei Geral de Proteção de Dados em *vacatio legis* no Brasil.

Ocorre que a indigitada MP, não previa diversos mecanismos de salvaguarda dos direitos dos titulares de dados, em total descompasso com as melhores práticas internacionalmente reconhecidas e com a lei que estava para entrar em vigor. Nesse sentido, a MP foi alvo de diversas ações direta de inconstitucionalidade (ADI) e o Supremo Tribunal Federal (STF) declarou a inconstitucionalidade liminar da referida MP, forte na argumentação da violação da autodeterminação do indivíduo em seu viés informacional.

Ato contínuo, o julgamento provocou o debate sobre a questão dos dados pessoais, especialmente dada a crise pandêmica, o que gerou uma vontade política na aprovação da Emenda Constitucional (EC) nº 115, de 2022 que

elencou como direito fundamental a proteção de dados pessoais, bem como serviu de incentivo para a entrada em vigor antecipada da Lei Geral de Proteção de Dados brasileira, n.º13.709/2018.

Pelo lado europeu, com a vanguarda que lhe é característica, a Corte Constitucional alemã já (*BvG)* desde os anos 1983, no clássico caso sobre o recenseamento da população alemã (*volkszählungsgesetz*) e, reconheceu a proteção de dados pessoais como uma derivação da autodeterminação informacional do indivíduo, ligado à sua dignidade (Mendes, 2020, p. 227).

Portanto, o BvG entendeu que não se poderia permitir que os dados pessoais recolhidos e tratados fossem utilizados com finalidade diversa daquela para qual a lei fora instituída, por violação do *nachteilverbot*, ou seja, a vedação de utilizar dados de forma diversa daquela declarada no momento da coleta.

Outrossim, a decisão deixou claro que a proteção a todo dado pessoal é importante, mesmo que aparentemente não tenha grande importância, pois somados a outros dados pessoais, possibilita uma análise global e completa, revestindo-se de significado.

Logo, a decisão consagrou a *autodeterminação informacional*, no sentido de que pertence à esfera individual de direitos a decisão própria sobre a utilização e tratamento dos dados pessoais, ou seja, o indivíduo tem o direito de controlar e decidir sobre seus dados como forma moderna de sua dignidade no viés da autonomia.

Segundo o saudoso doutrinador DONEDA (2019, p.169), a partir deste célebre julgamento, a *autodeterminação informacional* passou a ser concebida como um direito fundamental, na esteira do direito geral de personalidade, proporcionando ao indivíduo o controle de suas informações.

Alguns anos depois, com o fortalecimento da União Europeia, este entendimento fora consubstanciado em dois diplomas, quais sejam, a Diretiva 96/9/CE sobre proteção de dados pessoais e a Carta de Direitos Fundamentais da União Europeia, que consagrou a proteção de dados pessoais como direito fundamental no espaço da União, através do artigo 8.º, bem como o Regulamento Geral de Proteção de Dados (RGPD), que substituiu a diretiva retromencionada.

RODOTÁ já defendia a centralidade do direito fundamental à proteção de dados, como corolário da autodeterminação informacional, pugnando por certa autonomia em relação à proteção da privacidade, para ser uma garantia mais ampla e instrumental para a fruição de diversos outros direitos. Ou seja:

> A proteção de dados pessoais constitui não apenas um direito fundamental entre outros: é o mais expressivo da condição humana contemporânea. Relembrar isto a cada momento não é mera retórica, pois toda mudança que afeta a proteção de dados tem impacto sobre o grau de democracia que nós podemos experimentar (Rodotà, 2008, p.21).

Como bem nos ensina BRUNO BIONI (2021, p. 103), o direito à proteção de dados pessoais, além de proteger a personalidade, é imprescindível à liberdade, inovação e desenvolvimento de novas tecnologias:

> [...] a proteção de dados pessoais permite disciplinar a liberdade, a inovação e o desenvolvimento. E, em um cenário em que dados pessoais projetam a maneira como cada indivíduo é visto no mundo, permite também o exercício de direitos e da cidadania. Trata-se, hoje, do mais importante pilar do nosso contrato social (Bioni, 2021, p. 103).

Falar em proteção em face dos riscos da IA, portanto, é entender a centralidade do direito fundamental à proteção de dados pessoais.

3.2) Regime jurídico Brasil e EU e Interseções entre Proteção de Dados e Inteligência Artificial:

A IA é uma tecnologia estratégica, que oferece muitos benefícios para a sociedade, de forma que existem ganhos de eficiência e produtividade que podem fortalecer a competitividade dos negócios e melhorar o bem-estar das pessoas. Portanto, as questões de proteção de dados devem ser consideradas desde o início e monitoradas ao longo dos ciclos de vida dos sistemas de IA para garantir a conformidade com os direitos humanos e direitos fundamentais (Kuner *et.al*, 2018, p. 289).

Grandes desafios enfrentados pelos sistemas de IA como a opacidade, a complexidade, os preconceitos -ou enviesamentos-, imprevisibilidade, as decisões automatizadas e comportamentos parcialmente

autônomos de determinados sistemas de IA devem ser propriamente endereçados, a fim de garantir a compatibilidade destes sistemas com os direitos fundamentais e facilitar a aplicação das normas jurídicas.

Para CHRISTIAN TRONCOSO as pessoas têm o direito de saber como os seus dados estão sendo usados pelos sistemas de IA e, a tendência de IA e o acesso a dados pode ser uma das maneiras principais que podemos ter para regulamentar o uso de dados de maneira mais representativa. Então, claramente existe uma interseção da proteção de dados e inteligência artificial que levanta várias discussões, mas devemos reconhecer que inovação é importante, mas a proteção de dados e o cumprimento da Constituição também o é.

> Nós temos que estar conscientes dessa interseção e decidir, como uma sociedade, onde queremos delinear esses limiares, mas é importante, principalmente, que tenhamos leis transparentes e claras sobre como implementar essas regulamentações e que as empresas tenham orientações claras sobre o que é permitido e onde estão as limitações (Senado Federal, 2023, p. 136).

A LGPD, a RGDP e o PL 2338/23 - tal como se estabeleceu no AIA europeu-, adotaram uma linha de regulamentação baseadas em riscos (*risk-based approach*), estabelecendo-se diferenças de regime jurídico ou de carga regulatória com base no nível de riscos gerado pela atividade de tratamento de dados executada pelo agente de tratamento ou pelos sistemas de IA colocados no mercado.

Interseções, portanto, são naturais entre os dois temas e podem ser notadas na legislação, como os mecanismos de governança previstos, como o RIPD (relatório de impacto à proteção de dados) e o AIA (avaliação de impacto algorítmico), respectivamente, que auxiliam a promover estruturalmente nas organizações a conformidade ao regime de proteção de dados e às determinações da proposta de marco legal de IA*,* exatamente em linha com o que prega a doutrina do constitucionalismo digital, na medida em que advoga uma governação multinível- incluindo todos os atores privados e públicos no ônus de prover uma regulação-, como forma de endereçar respostas mais acuradas aos desafios do mundo digital, nomeadamente os ligados à IA (Sousa, 2022).

Nesta senda, a ANPD brasileira alertou que existem, notadamente, três campos de correspondência entre o PL n.º 2338/23 (PLIA) e a LGPD – *rectius* entre proteção de dados e IA- que devem ser destacados, dada a possibilidade de eventuais convergências, sobreposições e conflitos entre as duas legislações, são eles: i) direitos da pessoa afetada por sistema de IA e os direitos dos titulares; ii) a correlação entre sistemas de IA de alto risco e o tratamento de dados pessoais; e iii) mecanismos de governança (ANPD, 2023).

O primeiro ponto de contato entre as legislações de proteção de dados pessoais e de regulamentação de IA diz respeito aos direitos da pessoa afetada por sistemas de IA e o direito dos titulares de dados.

Para o PLIA, a informação tem papel fundamental na estruturação de uma governação bem sucedida, na medida em que a pessoa deve ter o máximo de informações claras e corretas, a fim de se autodeterminar de acordo com estas, em linha com o princípio da autodeterminação pessoal e informacional e autonomia humana.

Assim, nos termos do PLIA, à pessoa afetada devem ser asseguradas, previamente, informações claras e adequadas sobre uma série de aspectos, tais como: i) o caráter automatizado da interação com o sistema. ii) descrição geral do sistema, tipos de decisões, recomendações ou previsões que se destina a fazer e consequências de sua utilização para a pessoa; iii) identificação dos operadores do sistema de inteligência artificial e medidas de governança adotadas no desenvolvimento e emprego do sistema pela organização; iv) papel do sistema de inteligência artificial e dos humanos envolvidos no processo de tomada de decisão, previsão ou recomendação; v) categorias de dados pessoais utilizados no contexto do funcionamento do sistema de inteligência artificial; vi) medidas de segurança, de não-discriminação e de confiabilidade adotadas, incluindo acurácia, precisão e cobertura; e vii) outras informações definidas em regulamento

De mais, em consonância com o princípio da autodeterminação, o PLIA, confere os direitos de informação à pessoa, como i) explicação; ii) contestação; e iii) solicitar revisão, o que se torna demasiado importante também

para questões envolvendo decisões automatizadas, em função da potencialidade lesiva à direitos fundamentais dos utilizadores.

Logo, para o PLIA a pessoa afetada por sistema de inteligência artificial poderá solicitar explicação sobre a decisão, previsão ou recomendação, com informações a respeito dos critérios e dos procedimentos utilizados, assim como sobre os principais fatores que afetam tal previsão ou decisão específica, incluindo informações sobre: i) a racionalidade e a lógica do sistema; ii) o significado e as consequências previstas de tal decisão para a pessoa afetada; iii) o grau e o nível de contribuição do sistema de inteligência artificial para a tomada de decisões; iii) os dados processados e a sua fonte, os critérios para a tomada de decisão e, quando apropriado, a sua ponderação, aplicados à situação da pessoa afetada; iv) os mecanismos por meio dos quais a pessoa pode contestar a decisão; e v) a possibilidade de solicitar intervenção humana.

Em relação à explicabilidade das decisões de sistemas de IA, os agentes de IA que operem sistemas de alto risco também deverão adotar medidas técnicas para viabilizar a explicabilidade dos resultados dos sistemas de inteligência artificial e de medidas para disponibilizar aos operadores e potenciais impactados informações gerais sobre o funcionamento do modelo de inteligência artificial empregado, explicitando a lógica e os critérios relevantes para a produção de resultados, bem como, mediante requisição do interessado, disponibilizando informações adequadas que permitam a interpretação dos resultados concretamente produzidos, respeitado o sigilo industrial e comercial.

A supervisão humana de sistemas de inteligência artificial de alto risco buscará prevenir ou minimizar os riscos para direitos e liberdades das pessoas que possam decorrer de seu uso normal ou de seu uso em condições de utilização indevida razoavelmente previsíveis, viabilizando que as pessoas responsáveis pela supervisão humana possam, dentre outros direitos, intervir no funcionamento do sistema de inteligência artificial de alto risco ou interromper seu funcionamento e decidir, em qualquer situação específica, por não usar o sistema de inteligência artificial de alto risco ou ignorar, anular ou reverter seu resultado, segundo determina o parágrafo único do artigo 20 do PL n.º 2338/23.

Especificamente quanto ao direito de contestação e solicitação de revisão, há menção expressa à legislação de proteção de dados, em relação ao direito de solicitar a anonimização, bloqueio ou eliminação de dados desnecessários, excessivos ou tratados em desconformidade com a LGPD, num verdadeiro e expresso diálogo de fontes.

Ademais, pelo PLIA a pessoa afetada por sistema de inteligência artificial terá o direito de contestar e de solicitar a revisão de decisões, recomendações ou previsões geradas por tal sistema que produzam “efeitos jurídicos relevantes” ou “que impactem de maneira significativa seus interesses”, assegurando o direito de correção de dados incompletos, inexatos ou desatualizados utilizados por sistemas de inteligência artificial.

Importa ressaltar que quando a decisão, previsão ou recomendação de sistema de inteligência artificial produzir “efeitos jurídicos relevantes” ou que “impactem de maneira significativa os interesses da pessoa”, inclusive por meio da geração de perfis e da realização de inferências, esta poderá solicitar a intervenção ou revisão humana, sendo que em cenários nos quais as decisões, previsões ou recomendações geradas por sistemas de inteligência artificial tenham um impacto irreversível ou de difícil reversão ou envolvam decisões que possam gerar riscos à vida ou à integridade física de indivíduos, haverá envolvimento humano significativo no processo decisório e determinação humana final.

Esta determinação tem diálogo direto com o que ensina GIOVANNI DE GREGORIO (2023), para quem deve existir uma linha bem estabelecida de domínios onde as decisões automatizadas, realizadas amplamente por sistemas de IA, não poderão ter incidência sem a participação e supervisões humanas, a fim de evitar um *Rule of Tech*, ou seja, uma subordinação à normatização exercida por máquinas. Assim, para o PLIA, essa linha deveria ser colocada aquando da interferência no âmbito de direito da pessoa, desde que de forma significativa e relevante.

Similarmente ao direito de contestação e de solicitar revisão previstos no PLIA, a LGPD tem previsão do direito à solicitação de revisão de decisões

tomadas com base em tratamento automatizado. Ou seja, o titular dos dados tem direito a solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que “afetem seus interesses”, incluídas as decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade.

Importa destacar a visível sobreposição entre as legislações nesse sentido, o que poderá levar, inclusive, a alguma tensão interpretativa e regulatória por entidades designadas. Melhor dizendo, o titular dos dados tem direito a solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que “afetem seus interesses”, enquanto a previsão análoga no PL n.º2338/23 assegura que a pessoa afetada por sistema de inteligência artificial terá o direito de contestar e de solicitar a revisão de decisões, recomendações ou previsões geradas por tal sistema que produzam “efeitos jurídicos relevantes” ou “que impactem de maneira significativa seus interesses”, assegurando o direito de correção de dados incompletos, inexatos ou desatualizados.

Destarte, a LGPD e o PLIA trazem previsão de conceito jurídico indeterminado ao determinarem a revisão de decisões automatizadas caso haja afetação de interesse do titular e o direito de contestar e solicitar revisões quando o sistema de IA produzir efeitos jurídicos relevantes ou impactem de maneira significativa os interesses.

À evidência, a clarificação destes conceitos indeterminados ficará a cargo da pesquisa acadêmica, do entendimento dos órgãos administrativos reguladores e, em última instância, da jurisprudência do Poder Judiciário brasileiro. Mas não está errado concluir que a LGPD traz uma previsão mais abrangente e protetiva da pessoa afetada por decisões automatizadas, quando expressamente fala em afetação de interesses *lato sensu*. Claramente, afetação de interesses de maneira geral abrange uma maior gama de situações do que situações que causem efeitos jurídicos relevantes ou que impactem significativamente interesses do afetado.

Logo, em relação à interpretação destes dispositivos, a própria Lei de Introdução às Normas do Direito Brasileiro (LINDB) delineia balizas interpretativas que se aplicam neste particular.

Segundo determina a LINDB, nas esferas administrativa, controladora e judicial, não se decidirá com base em valores jurídicos abstratos sem que sejam consideradas as consequências práticas da decisão e a decisão administrativa, controladora ou judicial que estabelecer interpretação ou orientação nova sobre norma de conteúdo indeterminado, impondo novo dever ou novo condicionamento de direito, deverá prever regime de transição quando indispensável para que o novo dever ou condicionamento de direito seja cumprido de modo proporcional, equânime e eficiente e sem prejuízo aos interesses gerais.

Ao nosso entender a previsão do art. 20 da LGDP, por ser mais protetivo e amplo, deveria prevalecer e ser adotada em regra, inclusive para o caso de decisões automatizadas por sistemas de IA - em breve, será a esmagadora maioria de situações- justamente por ensejar consequências práticas mais benéficas e consentâneas com a tutela da pessoa e dos direitos fundamentais.

Uma interpretação a partir do princípio do *Antropocentrismo Digital* se faz necessária em razão da velocidade da inovação em IA, que está constantemente apresentando novas situações inéditas e a cautela aliada a maior proteção deverá ser a tônica para uma tutela adequada aos direitos fundamentais e do Estado de Direito.

Nesse sentido, a **“**Declaração Europeia sobre os Direitos e Princípios Digitais para a Década Digital”, assinada em 15 de dezembro de 2022, visando promover os valores europeus no âmbito da transformação digital**,** determina que as inovações devem dar prioridade às pessoas e compatibilizar o desenvolvimento econômico com os direitos fundamentais dos cidadãos e, portanto, é necessário criar uma agenda digital a ser implementada até 2030, onde se priorize a elaboração de normas que obedeçam as diretrizes

consubstanciadas no i) antropocentrismo; ii) liberdade de escolha; iii) segurança; iv) solidariedade e inclusão; v) participação e vi) sustentabilidade.

O princípio do *Antropocentrismo Digital*, portanto, significa que as tecnologias digitais devem proteger os direitos das pessoas, apoiar a democracia e assegurar que todos os intervenientes digitais agem de forma responsável e segura e, sob essa premissa, deverá ser interpretada e eventual colisão entre a LGPD e o PL n.º2338/23 sobre o grau de interferência de decisões automatizadas na esfera da pessoa atingida que autoriza o manejo de direitos de revisão e contestação.

Portanto, vez que todos estes direitos adrede apontados são direitos que se destinam à tutela das pessoas naturais afetadas por sistemas de IA, é de natural similaridade com alguns dispositivos da LGPD relativos aos direitos dos titulares, com o direito de acesso, erigido no art. 9º.

O direito de acesso, consagrado no art. 9º da LGPD, que também garante ao titular de dados o recebimento de informações relevantes sobre as operações de tratamento de seus dados pessoais. Tais informações devem ser disponibilizadas de forma clara, adequada e ostensiva acerca de, entre outros aspectos, a finalidade específica do tratamento, a forma e duração do tratamento. (ANPD, 2023)

Ou seja, o titular tem direito ao acesso facilitado às informações sobre o tratamento de seus dados, que deverão ser disponibilizadas de forma clara, adequada e ostensiva acerca de, entre outras características previstas em regulamentação para o atendimento do princípio do livre acesso: i) finalidade específica do tratamento; ii) forma e duração do tratamento, observados os segredos comercial e industrial; iii) identificação do controlador; iv) informações de contato do controlador; v) informações acerca do uso compartilhado de dados pelo controlador e a finalidade; vi) responsabilidades dos agentes que realizarão o tratamento; e vii) direitos do titular.

Por sua vez, o titular dos dados pessoais tem direito a obter do controlador, em relação aos dados do titular por ele tratados, a qualquer momento e mediante requisição: i) confirmação da existência de tratamento; ii)

acesso aos dados; iii) correção de dados incompletos, inexatos ou desatualizados; iv) anonimização, bloqueio ou eliminação de dados desnecessários, excessivos ou tratados em desconformidade com a LGPD; v) portabilidade dos dados; vi) eliminação dos dados pessoais tratados com o consentimento do titular; vii) informação das entidades públicas e privadas com as quais o controlador realizou uso compartilhado de dados; viii) informação sobre a possibilidade de não fornecer consentimento e sobre as consequências da negativa; e ix) revogação do consentimento.

Outrossim, tanto o PLIA (art. 3º, IV) quanto a LGPD (art. 6º, IX) mencionam o princípio da não discriminação. No caso do PL, esse princípio é ainda mencionado como fundamento (art. 2º, V) e direito (art. 5º, V), reforçando a proteção contra a discriminação, por meio de diversos instrumentos, como o direito à informação e compreensão, o direito à contestação, e em um direito específico de correção de vieses discriminatórios diretos, indiretos, ilegais ou abusivos, além das medidas de governança preventivas.

Entrementes, adotou as definições sobre discriminação direta e indireta – incorporando, assim, definições da Convenção Interamericana contra o Racismo, promulgada em 2022 –, tendo como ponto de atenção grupos (hiper)vulneráveis tanto para a qualificação do que venha ser um sistema de alto risco como para o reforço de determinados direitos.

Logo, o PLIA define discriminação, na linha da Convenção Interamericana contra o Racismo, como:

> Qualquer distinção, exclusão, restrição ou preferência, em qualquer área da vida pública ou privada, cujo propósito ou efeito seja anular ou restringir o reconhecimento, gozo ou exercício, em condições de igualdade, de um ou mais direitos ou liberdades previstas no ordenamento jurídico, em razão de características pessoais como origem geográfica, raça, cor ou etnia, gênero, orientação sexual, classe socioeconômica, idade, deficiência, religião ou opiniões políticas (Senado Federal, 2023).

E, ao seu turno, determina que a discriminação indireta é aquela que discriminação que:

> [...] ocorre quando normativa, prática ou critério aparentemente neutro tem a capacidade acarretar desvantagem para pessoas pertencentes a grupo específico, ou as coloquem em desvantagem, a menos que essa normativa, prática ou critério tenha algum objetivo ou justificativa razoável e legítima à luz do direito à igualdade e dos demais direitos fundamentais (Senado Federal, 2023).

Dessa forma, quando o PL nº2338/23 acaba por associar os efeitos discriminatórios aos usos ilegítimos e abusivos de dados pessoas sensíveis, definidos no art. 5º, II, da LGPD. Assim, fica claro que a identificação de efeitos discriminatórios de sistemas de inteligência artificial envolve, necessariamente, uma avaliação dos riscos associados ao tratamento de dados pessoais sensíveis (ANPD, 2023).

Nesse sentido, o noticiário está repleto de casos que se tornaram famosos pela discriminação gerada no uso de algoritmos, como o caso do programa COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions),* algoritmo originariamente concebido para gestão de penitenciárias, a partir de informações sobre detentos "críticos", que acabou tendo sido desvirtuada sua utilização para que o Poder Judiciário norte-americano utilizasse em julgamentos criminais, no sentido de avaliação do risco de reincidência de determinados réus, resultando em discriminações flagrantes, a partir de resultados produzidos sem qualquer respeito aos direitos constitucionais do processado, como *v.g.,* o aumento de pena em razão da possibilidade em abstrato de reincidência, a possibilidade de concessão de fiança ou não, além de mostrar-se tendencioso contra indivíduos latinos e negros[77].

Por conseguinte, direitos como a não discriminação está intimamente ligado à proteção de dados pessoais, vez que se cumprido o quadro-legal regulatório, há potencial chance de redução de discriminações diretas e indiretas realizadas por sistemas de IA. Todavia, para atingir esse nível de eficiência regulatória, mecanismos de governança e autorregulação são indispensáveis, especialmente num cenário de constitucionalismo digital, que propõe uma

[77] Para mais detalhamentos, conferir o caso de Eric L. Loomis vs. Wisconsin, disponível em https://harvardlawreview.org/2017/03/state-v-loomis/.

atuação multinível de esferas públicas, privadas, tanto nacional, quanto internacional.

Para a professora LAURA SCHERTEL MENDES a tutela jurídica dos dados pessoais pode auxiliar a combater a discriminação das informações oriundas da utilização das informações extraídas de banco de dados, buscando fornecer uma tutela mais rígida em caso de tratamento de dados sensíveis e de situações potencialmente discriminatórias

> A igualdade apresenta-se como um tema para essa disciplina, na medida em que a vigilância realizada por organismos privados e estatais, a partir de informações obtidas em bancos de dados, pode acarretar a seleção e a classificação dos indivíduos de modo a afetar a expressividade as suas oportunidades de vida em sociedade. Desse modo, a tutela jurídica dos dados pessoais pode auxiliar a combater a discriminação das informações oriundas da utilização das informações extraídas de banco de dados, buscando fornecer uma tutela mais rígida em caso de tratamento de dados sensíveis e de situações potencialmente discriminatórias. Essa proteção adquire maior importância quando se considera que, na sociedade atual, caracterizada pelas relações remotas, os dados pessoais acabem por se constituir na única forma de representação das pessoas perante as mais diversas organizações estatais e privadas, sendo determinantes para abrir ou fechar as portas das oportunidades e acessos'. Nesse sentido, entende-se fundamental a compreensão da disciplina de proteção de dados pessoais como meio de tutela da personalidade do cidadão, garantindo tanto a autonomia das escolhas como a sua proteção contra situações potencialmente discriminatórias (Mendes, 2014, p.28).

MENDES *et al.* (2021), relatam-nos que a expressão "discriminação algorítmica" compreende os seguintes cenários discriminatórios: "i) discriminação por erro estatístico; ii) discriminação pelo uso de dados sensíveis; iii) discriminação por generalizações injustas e iv) discriminação limitadora do exercício de direitos".

Entretanto, necessário explicar que, para os autores acima, o termo "discriminação algorítmica" deve englobar cenários que envolvam afirmações estatisticamente inconsistentes, quanto cenários em que as afirmações, embora

lógicas, traduzem discriminações por ausência de realce em determinada peculiaridade do indivíduo (Müller Dornelas, 2023).

Dessarte, os autores retromencionados trabalham com a noção de "generalizações" consistentes e inconsistentes, sendo que aquelas podem ser subdividas em universais (verdadeiras em 100% dos casos) e não universais (não se presta a descrever a totalidade de um grupo, mas sim uma característica compartilhada pela maioria desse grupo). Dessa maneira, a discriminação por erro estatístico significa que todo e qualquer erro que seja genuinamente assim, englobando dados erroneamente coletados, problemas nos códigos algorítmicos que resulte na contagem ou utilização errônea dos dados usados, representando, portanto, um erro de programação cometido por engenheiros de softwares ou cientistas de dados que desenvolveram a arquitetura do algoritmo.

Portanto, ao erigir ao *status* de princípio a determinação de não-discriminação, o PLIA e a LGPD dialogam com o princípio estruturante da igualdade, pedra angular do Estado de Direito, trazendo à conformação deste o tratamento de dados pessoais nos sistemas de IA.

Outro importante ponto de contato entre a temática de proteção de dados pessoais e regulação de sistemas de IA, especialmente extraído da LGPD e do PLIA, diz respeito ao tratamento de dados pessoais em sistemas de IA denominados de alto risco.

De forma análoga ao AIA, o PLIA subdivide a classificação do risco em excessivo ou alto, delimitando a regulação consoante os desdobramentos de risco da tecnologia e considerando os diferentes níveis de risco justificam diferentes níveis de restrições, garantias e salvaguardas proporcionais, haja vista a escolha por uma abordagem do tipo *risk-based*.

Assim, em algumas circunstâncias, o risco pode ser tão alto que, pela aplicação do princípio da precaução, pode ser necessário proibir o desenvolvimento ou uso de aplicativos dessas tecnologias, trazendo, inclusive, o conceito de riscos inaceitáveis ou utilizando de sistemas de *sandboxes*

regulatórios, a fim de perceber melhor os reflexos daquele(s) risco(s) em determinado(s) ambiente(s). (ANPD,2023)

Outra característica frequente nos sistemas considerados de alto risco é o tratamento volumoso de dados pessoais e de dados sensíveis – característica ínsita ao *big data*-. Ou seja, vários dos sistemas de inteligência artificial considerados de alto risco envolvem o tratamento de dados pessoais porque eles são projetados para tomar decisões automatizadas que podem ter um impacto significativo em direitos e interesses de indivíduos, como por exemplo, decisões relacionadas a crédito, emprego, segurança pública e saúde e, desta forma, acabam por oferecer maiores chances de violações de direitos dos usuários e pessoas.

Esses sistemas de IA de alto risco são treinados com algoritmos que se valem de bases de dados com grandes quantidades de dados pessoais, objetivando identificar padrões e tomar decisões mais precisas, como informações de identificação pessoal, histórico de compras, atividades realizadas on-line e até mesmo dados biométricos, par atingira o máximo de precisão e acurácia ao modelo algorítmico e sobressair mercadologicamente.

À evidência, salienta-se que dois dos critérios estabelecidos pelo PLIA para que a autoridade competente atualize as listas dos sistemas de IA de risco excessivo ou de alto risco envolvem expressamente a utilização de dados pessoais, quais sejam, i) alto nível de identificabilidade dos titulares dos dados; e ii) quando existirem expectativas razoáveis de confidencialidade ou de serem dados sensíveis do afetado nos sistemas de IA.

Vejamos:

> Art. 18. Caberá à autoridade competente atualizar a lista dos sistemas de inteligência artificial de risco excessivo ou de alto risco, identificando novas hipóteses, com base em, pelo menos, um dos seguintes critérios:
> [...]
> VIII – alto nível de identificabilidade dos titulares dos dados, incluindo o tratamento de dados genéticos e biométricos para efeitos de identificação única de uma pessoa singular, especialmente quando o tratamento inclui combinação, correspondência ou comparação de dados de várias fontes;

> IX – quando existirem expectativas razoáveis do afetado quanto ao uso de seus dados pessoais no sistema de inteligência artificial, em especial a expectativa de confidencialidade, como no tratamento de dados sigilosos ou sensíveis (Senado Federal, 2023).

Logo, existe uma preocupação regulatória especial com os efeitos desses sistemas de IA para os direitos fundamentais não apenas de seus usuários imediatos, mas de indivíduos e grupos que eventualmente estejam sujeitos às decisões desses sistemas, especialmente através de determinações e incentivos para uma autorregulação das empresas e *big tech.*

Filiamo-nos ao entendimento do pesquisador português SIMÃO SOUSA, para quem, partindo da premissa de uma nova esfera pública digital – como já discorrido no primeiro capítulo -, os serviços prestados, o que inclui àqueles operados por sistemas de IA, devem ser considerados como um serviço de interesse econômico geral e, por isso, devem ser regulados como se fossem verdadeiras utilidades públicas (Sousa, 2023).

A construção de uma governação digital multinível encontra suporte entre o constitucionalismo digital, mostrando apto a enfrentar problemas de ordem transnacional, partindo do princípio da dignidade da pessoa humana e dos princípios e valores constitucionais acordados de forma global, completando as legislações domésticas para questões que se situem num contexto global, de forma a estabelecer um sistema coeso na ordem interna e internacional a partir de uma denominador comum baseado na defesa do Estado de Direito (Sousa, 2023).

Assim sendo, o constitucionalismo digital oferece como resposta o *governance* multinível, onde todos os envolvidos devem assumir seus papeis e responsabilidades na cadeia de regulação da inovação, em especial as plataformas digitais, amplamente utilizadoras de dados pessoais e sistemas de IA:

> Com a imposição de limites decorrentes de uma lógica de governação multinível, proceder-se-á a uma alteração da arquitetura das próprias plataformas digitais da escolha do que aparece exposto ao utilizador, permitindo salvaguardar os processos democráticos pela limitação e maior controlo de

> disseminação informação falsa e dedicada a grupos específicos assentes em determinados perfis sociológicos, promovendo uma redução da personalização da informação e uma maior abertura à diversidade e pluralidade do conteúdo e informação apresentados permitindo uma maior fragmentação de conteúdo e melhoria de uma experiencia saudável e partilhada. (Sousa, 2023, p.127)

Portanto, quando se fala em governação eficaz de sistemas de IA – e da e plataformas digitais-, não se pode perder de vista a simbiose entre o público e o privado, a fim de estruturar um *framework* com modelos claros, transparentes e previamente determinados, composto por um sistema de camadas de atuação, onde há responsabilidade de entidades privadas e públicas.

E assim, arremata SIMÃO SOUSA:

> Em suma, cremos que a construção de um mecanismo de governação digital multinível partindo da cooperação entre Estados e comunidades de Direito, com as plataformas, assente na lógica e nos valores constitucionais, é aquela que melhor se posiciona para resolver um problema que existe e, ao dia de hoje, conhece apenas incipientes regulamentos sectoriais e de índole nacional, que sendo úteis localmente, fracassam espetacularmente no plano internacional (Sousa, 2023, p. 130).

Partindo da premissa de que uma regulação adequada deverá envolver sobretudo a participação de agentes privados, através de modelos de governação que destaquem a proteção e promoção dos direitos fundamentais, é imperioso comparar as medidas de *governance* entre os diplomas legais sobre proteção de dados pessoais e regulação de sistemas de IA, pelo que a análise comparada permite observar que as medidas de governança propostas pelo PL n.º2338/23 dialogam diretamente com diversos dispositivos da LGPD.

Logo, os princípios previstos no art. 6º, tais como transparência, segurança e não discriminação; os direitos dos titulares, em particular os relativos a decisões baseadas em tratamento automatizado de dados pessoais (art. 20); e o princípio da privacidade desde a concepção e por defeito (*privacy by design and default*), implícito no art. 46, §2º são evidências claras dessa interseção.

Outrossim, é patente certo semelhança entre as disposições sobre a instituição de programas de *governance* mediante códigos de conduta de agentes de IA (art. 30) com a disciplina da LGPD sobre regras de boas práticas e governança em proteção de dados, que estabelece, em seu art. 50, §2º, inciso I, o programa de governança em privacidade. Isto é, em ambos os casos se denota a escolha de política legislativa em favor da autorregulação regulada, a fim de promover nos agentes regulados atitude preventiva e de antecipação de riscos a direitos e liberdades fundamentais (ANPD, 2023).

Desse modo, o PLIA, através do comando do artigo n.º 19, determina que os agentes de inteligência artificial deverão estabelecer camadas de estruturas de governança e processos internos aptos a garantir a segurança dos sistemas e o atendimento dos direitos de pessoas afetadas, como:

> i) medidas de transparência quanto ao emprego de sistemas de inteligência artificial na interação com pessoas naturais, o que inclui o uso de interfaces ser humano-máquina adequadas e suficientemente claras e informativas; ii) transparência quanto às medidas de governança adotadas no desenvolvimento e emprego do sistema de inteligência artificial pela organização; iii) medidas de gestão de dados adequadas para a mitigação e prevenção de potenciais vieses discriminatórios; iv) legitimação do tratamento de dados conforme a legislação de proteção de dados, inclusive por meio da adoção de medidas de privacidade desde a concepção e por padrão e da adoção de técnicas que minimizem o uso de dados pessoais; v) adoção de parâmetros adequados de separação e organização dos dados para treinamento, teste e validação dos resultados do sistema; e vi) adoção de medidas adequadas de segurança da informação desde a concepção até a operação do sistema. (Senado Federal, 2023).

E, caso o sistema de IA seja classificado como de alto risco, para além das medidas supradescritas, adicionalmente, os agentes de IA que forneçam e operem estes sistemas, deverão observar, especialmente naquilo que concerne aos dados pessoais, medidas de gestão de dados para mitigar e prevenir vieses discriminatórios, incluindo:

> a) avaliação dos dados com medidas apropriadas de controle de vieses cognitivos humanos que possam afetar a coleta e organização dos dados e para evitar a geração de vieses por problemas na classificação, falhas ou falta de informação em

> relação a grupos afetados, falta de cobertura ou distorções em representatividade, conforme a aplicação pretendida, bem como medidas corretivas para evitar a incorporação de vieses sociais estruturais que possam ser perpetuados e ampliados pela tecnologia; e b) composição de equipe inclusiva responsável pela concepção e desenvolvimento do sistema, orientada pela busca da diversidade (Senado Federal, 2023).

Outro evidente paralelismo é extraído do instrumento de Avaliação de Impacto Algorítmico, com o Relatório de Impacto à Proteção de Dados Pessoais (RIPD). O Avaliação de Impacto Algorítmico parte de uma lógica que obriga os agentes de sistemas de inteligência artificial, sempre que o sistema for considerado como de alto risco pela avaliação preliminar, realizar estudo que compreende algumas etapas, a fim de permitir a descrição pormenorizada do sistema, identificar riscos e propor mecanismos de mitigação destes riscos. As etapas compreendem, pois, a i) preparação; ii) cognição do risco; iii) mitigação dos riscos encontrados; e iv) monitoramento.

Além de que, a avaliação de impacto algorítmico consistirá em processo iterativo contínuo, executado ao longo de todo o ciclo de vida dos sistemas de inteligência artificial de alto risco e, em atenção ao princípio da precaução, quando da utilização de sistemas de inteligência artificial que possam gerar impactos irreversíveis ou de difícil reversão, a avaliação de impacto algorítmico levará em consideração também as evidências incipientes, incompletas ou especulativas.

Ao seu turno, o RIPD, é definido no art. 5º, XVII como a documentação do controlador que contém a descrição dos processos de tratamento de dados pessoais que podem gerar riscos às liberdades civis e aos direitos fundamentais, bem como medidas, salvaguardas e mecanismos de mitigação de risco.

Por sua vez, o artigo n.º 38, parágrafo único, da LGPD, apresenta qual deve ser seu conteúdo mínimo a ser preenchido no relatório: i) a descrição dos tipos de dados coletados; ii) a metodologia utilizada para a coleta e para a garantia da segurança das informações; e iii) a análise do controlador com relação a medidas, salvaguardas e mecanismos de mitigação de risco adotados.

Embora o escopo de tratamento de um RIPD seja restrito a contextos que envolvam tratamento de dados pessoais e sua análise limitada à gestão de

riscos a liberdades e direitos fundamentais afetados em virtude deste tratamento, é indubitável a correlação entre as duas ferramentas, tanto no aspecto metodológico de sua elaboração, quanto no conteúdo, para os casos em que sistemas de IA tratem dados pessoais, como *v.g*. tecnologias de reconhecimento facial e aplicações de IA na área da saúde, que certamente exigirão a elaboração de AIA e RIPD. (ANPD, 2023).

À evidência, a interseção entre a proteção de dados e a inteligência artificial, especialmente quando se toma como parâmetro os diplomas em tela, mostra-se evidente e necessária. Assim sendo, para além de apontar para um enfoque à governação privada -ou até mesmo uma autorregulação regulada-, os direitos da pessoa afetada por sistemas de IA e os titulares de dados pessoais, bem como a correlação entre IA de alto risco e tratamento de dados pessoais surgem como os principais pontos de contato entre a LGDP e a PLIA.

Dessa forma, uma vez apontados os riscos gerais dos sistemas de IA e como o quadro legislativo de proteção de dados e de IA pretendem endereçar tais riscos, insta aprofundar a investigação quanto aos sistemas de IA generativa - que representam uma faceta irreversível da evolução da IA, com inúmero benefícios, mas igualmente com diversos desafios jurídicos, que pretendemos trazer a tona a partir da capítulo a seguir- e a proteção de dados, especialmente no contexto dos *Large Language Models* (LLMs), que possibilitam o funcionamento de IA Generativa.

4) Proteção de dados pessoais e Inteligência Artificial (IA) Generativa:

4.1) Introdução breve à conceitos, funcionamento e modelos de IA Generativa:

Desde o lançamento do ChatGPT no final de 2022, a IA generativa e os grandes modelos de linguagem (LLMs) conquistaram o mundo dada sua disruptividade. A nível técnico, podem ser distinguidos dos modelos de IA mais tradicionais de várias maneiras, em especial em razão de serem treinados em grandes quantidades de texto e geram linguagem como resultado, em oposição às pontuações ou rótulos na regressão ou classificação tradicional (Hacker *et al.*, 2023c)

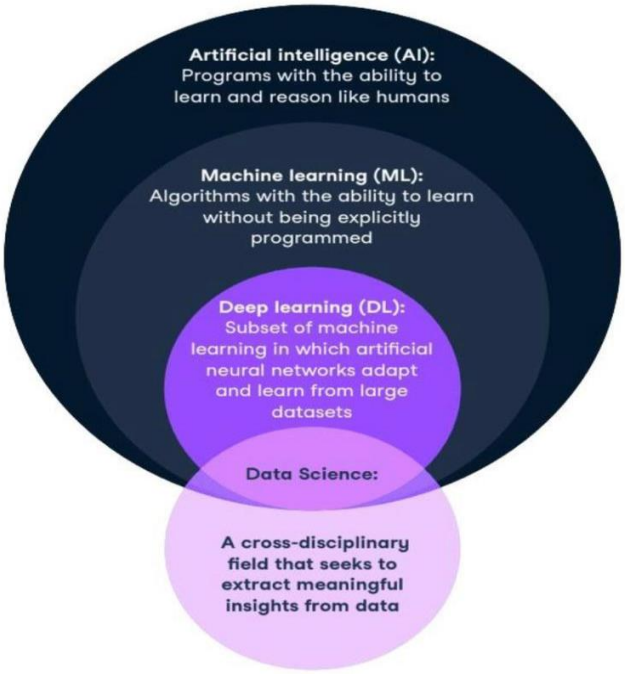
Assim sendo, compreende-se a inteligência artificial como a ciência da engenharia de fabricar máquinas inteligentes, especialmente programas de computador inteligentes, estando relacionado à tarefa de usar computadores para compreender a inteligência humana (McCarthy, 2007).

Nomeadamente existem dois tipos de inteligência artificial, quais sejam, de tipo forte e de tipo fraca. Melhor explicando, sistemas de IA de tipo forte ou *Strong AI,* caracterizam-se pela busca da consciência humana nas maquinas, ou seja, a IA forte é uma máquina com capacidade total habilidades cognitivas como os humanos, capazes de autoconsciência, aprendizagem, resolução de problemas e futuro planejamento. Atualmente, a IA forte permanece em grande parte teórica, existindo principalmente em romances ou filmes de ficção científica; no entanto, há uma previsão que os humanos consigam criá-lo com sucesso até o final deste século.

Em 2023, a OpenAI introduziu o GPT-4, um grande e versátil modelo de linguagem. Além de entradas de texto como seus antecessores, poderia usar imagens como entradas, identificar objetos e analisá-los para gerar respostas, sendo já considerado o GPT-4 como o inicial, embora incompleta, versão embrionária de um tipo de IA forte.

Recapitulando, dentro do recorte científico de estudo da inteligência artificial destacam-se os ramos de: i) Aprendizagem Automática (*Machine Learning*); ii) Aprendizagem Profunda (*Deep Learning*); Análise Preditiva (*Predictive Analytics*); e iv) processamento de Linguagem Natural (*Natural Language Processing*), sendo que todos esses subcampos da IA se valem fortemente de *big data*.

Observemos a imagem abaixo, que ilustra de forma didática as camadas da ciência afeta à inteligência artificial:

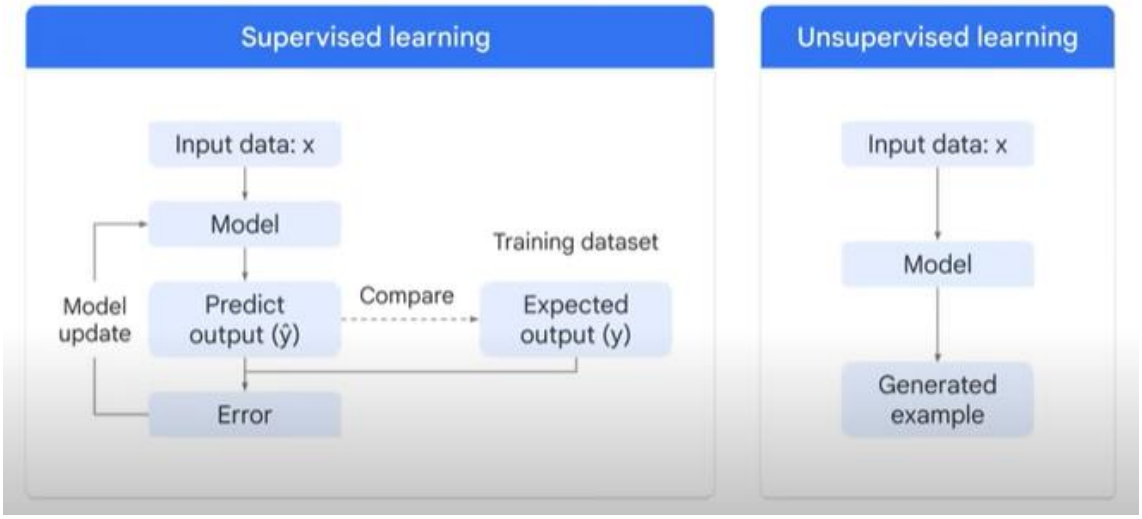


O ML era amplamente limitado a modelos preditivos, usados para observar e classificar padrões em conteúdo, o que se denomina de IA não-generativa. Por isso, encaixa-se dentro da subclassificação do tipo de IA fraca ou *weak AI*, mencionada no capítulo 2.

De mais, o ML opera numa lógica bipartida de modelos, quais sejam, supervisionados ou não-supervisionados, sendo a diferença fundamental entre os dois modelos é que nos de tipo “supervisionados” há a classificação ou rotulação dos dados utilizados, como um nome, número ou tipo e, portanto, implica dizer que o aprendizado é feito de exemplos passados para prever o futuro. Por sua vez, o modelo não-supervisionado os dados utilizados não são previamente rotulados e, por não se utilizar de dados não tratados previamente, o modelo tenta prever se os dados se encaixam naturalmente em algum grupo.

Vejamos o modelo em imagem disponibilizado pelo Google[78]:

[78] Observa-se que no modelo supervisionado, como existe a rotulagem dos dados utilizados, o sistema já sabe o que deseja ter em termos de resultado (*expected outcome*). Uma vez que o resultado alcançado com a previsão (*predict output*) não se encaixe com o desejado (expected) haverá erro. No sistema sem supervisão, ou seja, sem o etiquetamento prévio de dados, o modelo sugere um resultado que possa fazer sentido. Conf. https://cloud.google.com/learn/what-is-machine-learning?hl=pt-br. Acesso em 22 de Janeiro de 2024.



Ao seu turno, o *Deep Learning* (DL) ou aprendizagem profunda é um tipo de aprendizado de máquina que usa redes neurais artificiais para aprender com os dados e processar uma gama maior e mais complexa de padrões que o ML. Assim, as redes neurais artificiais são inspiradas no cérebro humano e podem ser usadas para resolver uma ampla variedade de problemas, incluindo reconhecimento de imagens, processamento de linguagem natural e reconhecimento de fala.

Algoritmos de aprendizagem profunda são normalmente treinados em grandes conjuntos de dados rotulados e estes aprendem a associar recursos nos dados aos rótulos corretos, tal como uma tarefa de reconhecimento de imagem, o algoritmo pode aprender a associar certas características de uma imagem ao rótulo correto.[79] Depois que um algoritmo de DL for treinado, ele poderá ser usado para fazer previsões sobre novos dados, como *v.g* um algoritmo de aprendizagem profunda que foi treinado para reconhecer imagens de cães pode ser usado para identificar cães em novas imagens.

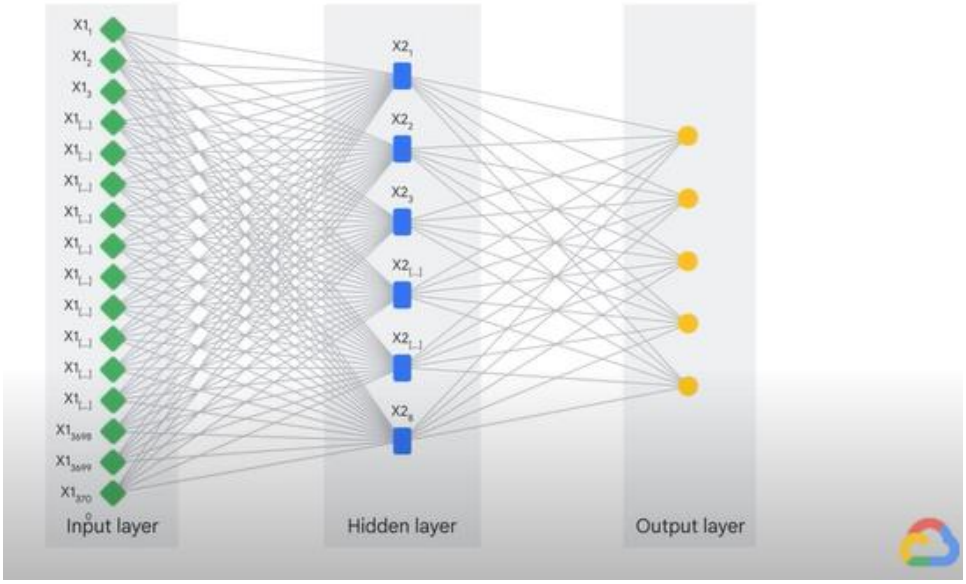
O *Deep Learning* funciona usando redes neurais artificiais para aprender com os dados e estas são compostas por camadas de *nodes* ou "nós" interconectados, e cada destes é responsável por aprender uma característica específica dos dados. Com base no nosso exemplo anterior com imagens –

[79] Conf. https://cloud.google.com/discover/what-is-deep-learning. Acesso em 22 de janeiro de 2024.

numa rede de reconhecimento de imagens, a primeira camada de *nodes* pode aprender a identificar arestas, a segunda camada pode aprender a identificar formas e a terceira camada pode aprender a identificar objetos[80].

À medida que a rede aprende, os pesos nas conexões entre os *nodes* são ajustados para que a rede possa classificar melhor os dados. Esse processo é chamado de treinamento e pode ser realizado por meio de diversas técnicas, como aprendizagem supervisionada, aprendizagem não supervisionada e aprendizagem semi-supervisionada.

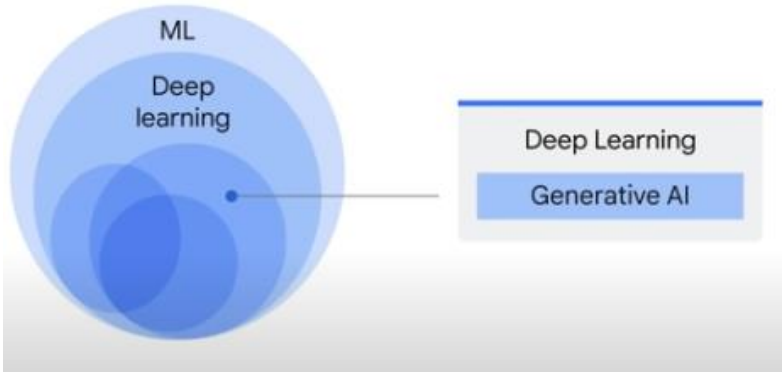
Dessa forma, DL funciona como um cérebro humano, interconectando os nodes (neurônios) que podem aprender a realizar determinadas tarefas processando dados e fazendo previsões, a partir de várias camadas de nodes que os permite aprender tarefas mais complexas que o tradicional ML. Vejamos[81]:



Assim, ao combinar aprendizagem supervisionada com aprendizagem não-supervisionada, a rede neural de aprendizagem é treinada com uma pequena quantidade de “dados etiquetados” (*label data*) em relação aos “dados não-rotulados” (*unlabel data*), de sorte que os “dados etiquetados” permite que a rede neural tenha referências básicas para aprendizado, enquanto os “dados sem rotulação” fazem a função de generalizar para novos exemplos

[80] Ibidem.
[81] Ibid.

nos resultados ou *outcomes*, chegando ao que convencionou-se denominar de "inteligência artificial generativa". Observemos a imagem ilustrativa abaixo[82]:



A *inteligência artificial generativa*[83], pois, refere-se a um subcampo da inteligência artificial, mais precisamente do *Deep Learning*, que se concentra na criação de sistemas capazes de gerar conteúdo original e criativo, como texto, imagens, música, vídeos e muito mais. Esses sistemas são projetados para aprender padrões a partir de grandes conjuntos de dados e usar esse conhecimento para produzir novos conteúdos de forma autônoma.

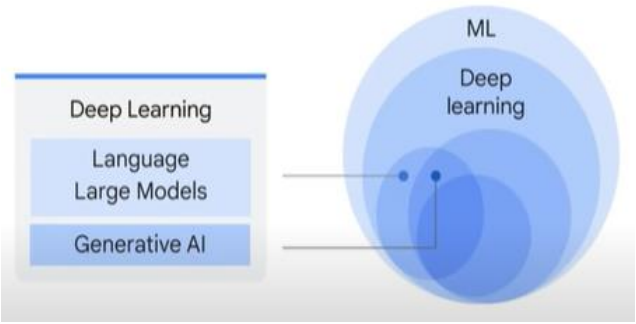
Existem diversas técnicas de IA generativa, como as Redes Neurais Generativas (GANs); as Redes Neurais Recorrentes (RNNs); Generative Pre-trained Transformers; e Redes Neurais Convolucionais (CNNs), conforme já expusemos em capítulo anterior[84].

A IA generativa foi um avanço, pois em vez de simplesmente entender e classificar o padrão, o aprendizado de máquina agora é capaz de criar novos

[82] Ibid.
[83] Para além do modelo de *Generative AI* existe o *Discriminative AI*, que se utiliza de classificação e predição a partir do uso de dados etiquetados ou rotulados somente, aprendendo da interação entre os *nodes* e estes dados.
[84] As GANs são um tipo popular de modelo de IA generativa que consistem em duas redes neurais, uma geradora e uma discriminadora, que competem entre si. A rede geradora cria dados sintéticos, enquanto a discriminadora tenta distinguir entre dados reais e sintéticos. Esse processo iterativo leva à geração de dados cada vez mais realistas; as RNNs são usadas para gerar sequências de dados, como texto ou música. Elas são capazes de manter informações contextuais ao longo do tempo, tornando-as adequadas para tarefas de geração de linguagem natural; por sua vez, *transformers* são os modelos de linguagem baseados em transformação de dados, como o ChatGPT (*Generative Pre-trained Transformer*), são amplamente utilizados para tarefas de geração de texto. Eles são treinados em grandes quantidades de texto para aprender a estrutura e o estilo da linguagem, permitindo a geração de texto coerente e contextualmente relevante; as CNNs são frequentemente usadas na geração de imagens. Elas podem aprender padrões visuais e criar novas imagens com base nesses padrões.

padrões sem *inputs* humanos. Portanto, com a compreensão adquirida, a IA generativa se torna capaz de gerar e criar novos dados como textos, imagens e áudios, semelhantemente aos utilizados na formação inicial do padrão, sem a necessária intervenção prévia humana na classificação e padronização dos dados. O avanço deste tipo de modelo levou a tecnologia para a fronteira do conhecimento, dando um largo passo para uma possível consciência ou um modelo de IA Forte, o que traz benefícios e prejuízos.

Assim, a IA generativa usa um modelo de DL, mais precisamente os *Large Language Models* (LLMs) para aprender os padrões e as relações em um conjunto de dados de conteúdo criado por humanos; em seguida, ele usa os padrões aprendidos para gerar novo conteúdo. Em suma, *Large Language Models* (LLMs) -ou grandes modelos de linguagem- são um subcampo do *Deep Learning* (DL) que consiste, basicamente, no pré-treinamento e, depois, ajustados com precisão (fine-tuned) para fins específicos e possui uma interseção natural com a inteligência artificial generativa[85], como podemos perceber da ilustração elaborada pelo Google, logo abaixo[86]:



Ou seja, os LLMs são grandes modelos de linguagem que consistem tipicamente em uma rede neural profunda com muitos parâmetros, treinada com grandes quantidades de texto não rotulado usando aprendizado de máquina semi-supervisionado. (Ramos *apud* Digampietri, 2023). O avanço dos LLMs provocou uma disrupção no campo do processamento de linguagem natural NLP

[85] Amazon Web Services, “What are Large Language Models (LLM)?”, https://aws.amazon.com/what-is/large-language-model/. Acesso em 23 de janeiro de 2024
[86] Conf. https://cloud.google.com/use-cases/generative-ai?hl=pt-BR acesso em 22 de janeiro de 2024.

em função da dinâmica de aprendizado de máquina que permite que um computador entenda, analise e simule a linguagem humana.

Frequentemente, os LLMs são marcados por seu escopo mais amplo e maior autonomia na extração de padrões em grandes conjuntos de dados. Em particular, a capacidade dos LLMs de escalabilidade geral permite-lhes gerar conteúdo processando uma gama variada de entradas de vários domínios, sendo que muitos dos LLMs são multimodais, o que significa que podem processar e produzir vários tipos de formatos de dados simultaneamente, como o ocorre no GPT-4, que pode lidar com entradas de texto, imagem e áudio simultaneamente para gerar *outcomes*.

Outrossim, pode-se dividir o conceito em três grandes partes, quais sejam, “large”; “general-purpose”; e “pre-trained and fine-tuned”. O termo “large” refere-se i) à enorme quantidade de dados utilizados no modelo, chegando à casa dos “Petabyte”[87]; e ii) massivo número de parâmetros[88], que são as memórias e o conhecimento aprendido pelo modelo de treinamento que serão utilizados para cumprir a tarefa do modelo. Ao seu turno, o termo “general-purpose” indica que que a concepção do modelo de LLMs é suficiente para resolver problemas corriqueiros e mais comuns, em função da generalidade da linguagem humana e da restrição de recursos, que implica reconhecer que poucas empresas no mundo tem condições de financiar esse treinamento pelo número de dados que são necessários e os custos; por fim, “pre-trained and fine-tuned” dizem respeito a segunda etapa consistente em refinar e dar acurácia par tarefas menores e mais específicas, após o treinamento geral todavia, a partir de uma diminuta quantidade dados.

Ou seja, os LLMs são treinados comumente para exercer tarefas gerais como a classificação de textos e sumarização de documentos; responder perguntas e criar textos e imagens e, a partir destas funções, o ajuste fino poderá

[87] Um petabyte é uma medida de memória ou capacidade de armazenamento de dados igual a 2 elevado à 50ª potência de bytes. Existem 1.024 terabytes (TB) em um petabyte e aproximadamente 1.024 PB constituem um exabyte. Disponível em https://www.techtarget.com/searchstorage/definition/petabyte. Acessado em 23 de janeiro de 2024.

[88] A Google lançou em 2022 o *Pathways Language Model* (PaLM) que possui um parâmetro de 540 bilhões. Disponível em https://blog.research.google/2022/04/pathways-language-model-palm-scaling-to.html. Acesso em 23 de janeiro de 2024.

ser realizado para acurar o modelo à determinados clientes ou mercados, como por exemplo o varejo, financeiro, mídia e entretenimento, fornecendo um serviço único e desejado por grandes empresas, em função da precisão do resultado e da otimização de tempo dispendido nas tarefas, tudo isso usando os dados como a principal matéria-prima.

Seus benefícios são reconhecidamente i) utilização de um mesmo modelo para diferentes tarefas, em função do volume (*petabyte*) de dados utilizados, que geram bilhões de possibilidades e parâmetros ao modelo, o que permite utilizá-lo para diversas funções; ii) o ajuste fino para adequação mais precisa à determinado serviço ou produto, o que demanda uma quantidade de dados bem menor; e iii) a retroalimentação do sistema conjugado de grandes dados para propósito geral, com dados menores de propósito específico, que faz com que o modelo aprenda cada vez mais e, assim, necessite de menos dados para ter a mesma precisão de antes.

Grandes modelos de linguagem servem como modelos básicos, fornecendo uma base para uma ampla gama de tarefas de processamento de linguagem natural (PNL) e a IA generativa pode abranger uma série de tarefas além da geração de linguagem, incluindo geração de imagens e vídeos, composição musical e muito mais, vez que grandes modelos de linguagem, como uma aplicação específica de IA generativa, são projetados especificamente para tarefas que giram em torno da geração e compreensão de linguagem natural, com a massiva utilização de conjuntos de dados para aprender padrões e relações entre palavras e frases.

No entanto, embora os LLMs avançados geralmente tenham um bom desempenho em um amplo espectro de tarefas, isso traz resultados altamente imprevisíveis, mesmo para seus criadores, levantando preocupações sobre a legalidade e a precisão dos textos gerados pelo LLMs (Hacker *et al.* 2024 *apud* Ganguli *et al*. 2022). Consequentemente, a tutela dos dados pessoais – e de dados e geral- torna-se imprescindível para minimizar os efeitos adversos destes modelos.

4.2) Detalhando impactos e aplicações da IA Generativa: alucinações, dados sintéticos e *modelos de treinamento*)

Conforme se percebe da estrutura de funcionamento da ciência ligada à inteligência artificial - que compreende, dentre outros, o *Machine Learning*, *Deep Learning*, *Large Language Models*, IA generativa– os dados ocupam a centralidade do sistema. Assim, pode-se dizer que, primeiramente, os dados são os maiores impactados pelo incorreto e ilícito uso das novas tecnologias, mas se tornam, ao mesmo tempo, a esperança de uma regulação eficaz, caso haja o devido compliance ao quadro legal de proteção de dados pessoais, *ex vi* da LGPD e do RGPD.

Grande preocupação diz respeito a coleta e tratamento destes dados utilizados nos LLMs. Assim, quanto a origem desta enorme quantidade de dados não existe muita transparência, pelo que desenvolvedores de IA não divulgam os detalhes exatos de seus conjuntos de dados de treinamento. Para IA generativa, a maior parte dos dados de treinamento são “extraídos” de páginas da Web disponíveis publicamente antes de serem reembalados e vendidos ou, em alguns casos, disponibilizados gratuitamente para desenvolvedores de IA.

Alguns desenvolvedores de IA contam com grandes conjuntos de dados populares, como “Colossal Clean Crawled Corpus” (C4) e “Common Crawl”, que são acumulados por meio de rastreamento da web (ou seja, software que navega sistematicamente em sites públicos da Internet e coleta informações de cada página da web disponível). Da mesma forma, os geradores de imagens de IA são normalmente treinados num conjunto de dados chamado LAION, que contém bilhões de imagens extraídas de sites da Internet e as suas descrições de texto. Algumas empresas também podem utilizar conjuntos de dados de sua propriedade para formação de modelos (Busch, 2023).

De mais a mais, os conjuntos de dados de IA generativa podem incluir informações publicadas em sites da Internet disponíveis publicamente e conteúdo sensível e protegido por direitos autorais. Eles também podem incluir conteúdo disponível publicamente que seja errôneo, pornográfico ou potencialmente prejudicial[89].

[89] Em uma investigação de 2023, o Washington Post e o Allen Institut for AI analisaram os sites extraídos para o conjunto de dados C4, que é usado por desenvolvedores de IA, incluindo

Logo, partindo da premissa que não há a transparência necessária quanto a submissão destes dados ao correto tratamento, diversos riscos são colocados em debate, como aqueles elencados no *report* "*Safety and Security Risks of Generative Artificial Intelligence to 2025"* do Reino Unido. O relatório supra afirma que é altamente provável que a IA generativa amplifique os riscos existentes do que crie riscos totalmente novos, mas aumentará drasticamente a velocidade e a escala de algumas ameaças, como o aumento de vulnerabilidades digitais; a erosão da confiança na informação; Influência política e social; uso inseguro e uso indevido, dentre outros.

Portanto, a IA generativa poderá levar à poluição do ecossistema de informação pública com *bots* hiper-realistas e meios de comunicação sintéticos (deepfakes) que influenciam o debate social e refletem preconceitos sociais pré-existentes. Este risco inclui a criação de notícias falsas, a desinformação personalizada, a manipulação dos mercados financeiros e o enfraquecimento do sistema de justiça criminal.

À evidência, as ferramentas generativas de IA são capazes de persuadir os seres humanos sobre questões políticas e podem ser utilizadas para aumentar a escala, o poder de persuasão e a frequência da desinformação. Ainda, a integração de IA generativa em sistemas e infraestruturas críticas poderá desencadear sistemas tendenciosos e discriminatórios ou no comprometimento da tomada de decisões humanas devido a uma segurança da informação deficiente e a processos algorítmicos opacos, por exemplo, "alucinações", ou seja, onde que sistemas de IA e LLMs, produzem informações incorretas, fictícias ou que não estão fundamentadas nos dados de *input*, criando literalmente conteúdo sem sentido ou falso.

Portanto, muito dos problemas da IA generativa resumem-se a dados insuficientes ou não rotulação adequada, gerando resultados sem precisão e ilegítimos, cujas consequências são desastrosas em todos os

Google, Facebook e OpenAI. A investigação descobriu que o conjunto de dados C4 incluía sites com conteúdo protegido por direitos autorais como bem como informações potencialmente confidenciais, como registros estaduais de recenseamento eleitoral.

aspectos da sociedade, como discriminações odiosas, desinformação, fake news, polarização política, religiosa, de gênero e racial.

Desta maneira, os dados sintéticos se apresentam como uma possível solução para resolver os problemas de dados insuficientes e imprecisos, vez que produzem dados artificiais “do zero” ou usando técnicas avançadas de manipulação de dados para produzir exemplos de treinamento novos e diversos, minimizando resultados enviesados e sem precisão, por um custo menor do que a tradicional rotulação manual de dados (Nikolenko, 2022).

Destarte, os dados de formação são, de uma perspectiva técnica, de importância fundamental para o desenvolvimento de aplicações de IA. Os dados de treinamento são a base para ambos ambientes de aprendizagem supervisionada e simulação na área de aprendizagem por reforço. Essas duas técnicas, por sua vez, são a base para a maioria das aplicações de IA atualmente em uso, desde reconhecimento facial automatizado, pontuação de crédito e recrutamento de IA para desempenho supra-humano de Agentes de IA em vários jogos complexos (Hacker *et al*. 2024).

Na aprendizagem supervisionada, o modelo algorítmico é calibrado combinando previsões com resultados (supostamente) corretos já contidos nos dados de treinamento. Na aprendizagem por reforço, por outro lado, a IA desenvolve uma estratégia que é determinada com base em um ambiente de aprendizagem que consiste em dados que enviam sinais de recompensa (*feedback*) ao modelo. (Hacker *et al*. 2024).

Dada as promessas e riscos associados à IA, os dados de treinamento representam, portanto, um elemento regulatório fundamental problema para a *sociedade algorítmica*. Desta maneira, o correto uso e tratamento de dados em sua base mostra-se vital para manutenção de sistemas de IA, especialmente os generativos, confiáveis e com os riscos minimizados.

Como adverte-nos PHILIPP HACKER, os dados de treinamento ainda representam comparativamente terra incógnita no mundo jurídico e sua importância central para técnicas de aprendizado de máquina, no entanto,

sugere que, ao contrário de uma visão generalizada, não se trata tanto de uma regulamentação de algoritmos, mas de uma regulamentação de dados necessários – em particular, dos dados de treinamento de IA. (Hacker, 2021, p. 6).

De mais, aponta o pesquisador para três riscos centrais e interligados e na forma como são abordados na legislação da UE ou dos Estados-Membros (harmonizados):i) riscos para a qualidade dos dados; ii) riscos de discriminação; e iii) riscos de inovação.

Para HACKER (2021) Em relação aos riscos de qualidade, os dados são fundamentais para o aprendizado de máquina e, por isso, eles têm implicações diretas para técnicas de aprendizagem supervisionada porque dados de treinamento objetivamente incorretos (normalmente) levam a previsões incorretas do modelo “alucinações”. No entanto, a qualidade dos dados não se limita à correção objetiva, mas também deve incluir, por exemplo, a atualidade e a representatividade dos dados, para prevenir também os vieses.

Dessarte, consoante as lições do pesquisador DIEGO MACHADO (2023), no âmbito da proteção de dados, a qualidade dos dados remonta desde as primeiras leis e documentos dos idos da década de 1970, à noção de exatidão (*accuracy*) dos dados objeto do tratamento, de sorte a receber a nomenclatura de princípio da exatidão (*data accuracy principle*), sendo previsto na RGPD, através do artigo 5.º, 1, c, onde determina que os agentes de tratamento devem assegurar com razoável segurança que adotarão as medidas necessárias para tratar os dados exatos e atualizado e, igualmente, previsto como princípio na LGPD, no inciso VI do artigo 6.º.

Entrementes, o uso do princípio da qualidade de dados é primordial para tutela dos dados em sistemas de IA generativa, especificamente em relação aos dados de treinamento para LLMs, pois é deste princípio que derivam direitos como bloqueio, apagamento e retificação de dados.

Quanto aos riscos discriminatórios, os dados de treinamento também são uma importante fonte de discriminação algorítmica, comprovado através dos

famosos casos reais nas áreas de reconhecimento facial, recrutamento por sistemas IA e publicidade personalizada.

Os riscos de discriminação estão parcialmente ligados ou podem ser uma consequência dos riscos de qualidade dos dados se na medida em que a qualidade dos dados para um determinado grupo protegido, especialmente minorias, for em média, afetado negativamente. Contudo, esta ligação não existe necessariamente, vez que podem surgir riscos de discriminação independentemente dos riscos de qualidade.

Ou seja, mesmo que a qualidade dos dados seja a mesma em relação aos diferentes grupos protegidos, a falta de equilíbrio do grupo em um conjunto de dados (por exemplo, a sub-representação de um grupo protegido, o chamado viés de amostragem) pode levar a distorções sistemáticas e discriminação (Hacker, 2021).

No entanto, ainda alerta HACKER (2021) que se deve reconhecer que as decisões tomadas por seres humanos também podem ser guiadas em grande medida por preconceitos conscientes ou inconscientes, conhecidamente denominado de preconceitos estruturais.

Por isso, uma preocupação significativa gira em torno da amplificação involuntária de preconceitos e a promoção de discriminação nos sistemas de IA. Esses sistemas dependem em conjuntos de dados extensos que podem incorporar inerentemente preconceitos derivados de escolhas humanas passadas. Consequentemente, a IA pode reforçar questões de gênero, raça ou preconceitos socioeconômicos, resultando em discriminações consequências, nomeadamente em domínios como o emprego, empréstimos e o sistema de justiça.

No que tange aos riscos de inovação, num ambiente técnico e dinâmico como o da IA, eles são divididos em duas dimensões, quais seja, riscos de bloqueio e riscos de regulamentação excessiva.

Primeiramente, fala-se em “risco de bloqueio” para a inovação, ou seja, isso ocorre porque os dados podem estar sujeitos a direitos de propriedade

intelectual ou podem ser protegidos por leis de proteção de dados; isso, por sua vez, faz com que seu uso como dados de treinamento consideravelmente mais difícil (Hacker, 2021).

Em segundo lugar, existe um risco global de regulamentação excessiva, que pode inibir indevidamente a desenvolvimento da IA devido a custos significativos ou mesmo proibitivos para os destinatários. Trata-se, no entanto, antes de mais, de uma questão de calibrar os respectivos ônus, que deverá ser considerado, a seguir, nos requisitos legais individuais abordar os riscos que acabamos de mencionar. (Hacker, 2021).

Outrossim, os dados sintéticos se apresentam como uma possível solução para responder aos problemas de dados insuficientes e imprecisos por reduzir resultados enviesados e sem precisão, por um custo menor do que a tradicional rotulação manual de dados .

Desta forma, importante destacar que os dados criados pela IA generativa, não criados por humanos, mas que imitam dados do mundo real, são chamados de dados sintéticos. Estes são produzidos por algoritmos e simulações de computação baseados em tecnologias de inteligência artificial generativa. De sorte que um conjunto de dados sintéticos tem as mesmas propriedades matemáticas dos dados reais nos quais ele se baseia, mas não contém as mesmas informações.

Por isso, inovações recentes em IA tornaram a geração de dados sintéticos eficiente e rápida, mas também aumentaram sua importância em questões regulatórias de dados[90]. Como vantagens da utilização de dados sintéticos, afora a questão de redução de custos já falado acima, listamos o i) aumento da proteção de dados e privacidade, ao diminuir consideravelmente o risco de exposições de informações pessoais sensíveis que podem ser encontradas em dados reais; ii) disponibilidade de dados, quando a oferta de

[90] Conf. https://aws.amazon.com/pt/what-is/synthetic-data/#:~:text=Dados%20sint%C3%A9ticos%20s%C3%A3o%20dados%20n%C3%A3o,tecnologias%20de%20intelig%C3%AAncia%20artificial%20generativa. Acesso em 23 de janeiro de 2024.

dados em determinada área for escassa; iii) viés, na medida em que ao criar o dado é possível controlar ou mitigar vieses.

Todavia, embora ofereçam vantagens, os dados sintéticos trazem preocupações, tais como: i) viés sintético, que corresponde enviesamento do próprio dado sintético, contaminado pelo banco de dados reais ou do algoritmo; ii) substituição de dados reais, no sentido de que a dependência excessiva de dados sintéticos poderá desaguar numa menor coleta de dados reais de alta qualidade, diminuindo a acuracidade, especialmente para nuances e detalhes, no treinamento do sistema para gerar dados sintéticos; iii) fidelidade aos dados reais, ou seja, dados sintéticos precisam ser suficientemente fiéis aos dados reais para serem úteis, se não capturarem com precisão as distribuições, relações e padrões presentes nos dados reais, os modelos treinados com esses dados podem não se comportar adequadamente em cenários do mundo real; iv) de mais, há sempre a possibilidade de incerteza quanto a qualidade dos dados utilizados, pelo que a qualidade dos dados sintéticos pode variar dependendo das técnicas de geração utilizadas e das suposições subjacentes[91].

Em todos os casos acima, especialmente acerca de enviesamento de dados, notadamente em razão de replicação de padrões discriminatórios enraizados culturalmente, esbarramos novamente naquilo que Hacker chama de risco de qualidade e de discriminação, mais precisamente na precisão e qualidade dos dados de treinamento (Hacker, 2021).

Estas preocupações podem ser particularmente pertinentes para sistemas de IA generativos utilizados em interações ou serviços que normalmente resultam na divulgação de informações sensíveis, tais como aconselhamento, cuidados de saúde terapêuticos, serviços jurídicos ou financeiros.

Conforme se presume no relatório acima citado, os dados sintéticos gerados por a IA generativa, até 2026, poderão abranger uma grande proporção

[91] Conf. https://www.iaresponsavel.com.br/2023/08/31/o-que-sao-dados-sinteticos/. Acesso em 23 de janeiro de 2024.

de conteúdo on-line e correm o risco de minar a confiança do público no governo, ao mesmo tempo que aumentam a polarização e o extremismo.

Por isso, estes riscos oriundos dos sistemas de IA generativa já suscitaram o debate na sociedade e, inclusive, medidas concretas de proteção de dados já foram tomadas ao redor do mundo, especialmente no âmbito europeu – dada a magnitude e maturidade do tema no espaço da UE-, onde a Autoridade Nacional de Proteção de Dados Pessoais da Itália suspendeu em março de 2023 o funcionamento do ChatGPT, operado pela empresa OpenAI, em todo o território italiano [92], numa nítida homenagem ao princípio da precaução digital, pelo qual caso haja dúvidas significativas sobre os efeitos nocivos de determinada tecnologia, deve-se optar pela cautela, até que se perceba melhor como mitigar ou atenuar tais consequências.

Segundo a autoridade daquele país, a falta de informação aos utilizadores e a todos os titulares de dados cujos dados são recolhidos pela OpenAI, mas sobretudo a ausência de uma base legal que justifique a recolha e armazenamento massivo de dados pessoais, de forma a "treinar" os algoritmos subjacentes ao funcionamento da plataforma foram determinantes para a decisão de suspensão (Garante per la Protezione dei Dati Personali, 2023).

Outrossim, como evidenciado pelas verificações realizadas, as informações fornecidas pelo ChatGPT nem sempre correspondem aos dados reais, resultando em processamento inexato de dados pessoais. Por fim, apesar de – de acordo com os termos publicados pela OpenAI – o serviço se destinar a maiores de 13 anos, a Autoridade salienta que a ausência de qualquer filtro para verificar a idade dos utilizadores expõe os menores a respostas absolutamente inadequadas no que diz respeito ao seu grau de desenvolvimento e autoconhecimento. (Garante per la Protezione dei Dati Personali, 2023).

Tal intervenção estatal da autoridade italiana, também, para além da perspectiva de proteção em vista dos riscos de qualidade e de discriminação,

[92] Conf. www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/9870847. Acesso em 23 de janeiro de 2024.

configura demonstração do risco de inovação, especificamente no bloqueio pautado na legislação de proteção de dados pessoais.

No mesmo sentido, a *Agencia Española de Protección de Datos* iniciou de ofício um processo de investigação preliminar contra a empresa norte-americana *OpenAI,* proprietária do serviço ChatGPT, por possível incumprimento dos regulamentos, solicitando ao Comitê Europeu para a Proteção de Dados (CEPD) – do qual a AEPD é membro juntamente com outras autoridades de proteção de dados do EEE – que incluísse o serviço ChatGPT como uma questão a abordar na sua reunião plenária, considerando que o tratamento global pode ter um impacto significativo nos direitos das pessoas exigem ações harmonizadas e coordenadas a nível europeu em aplicação do Regulamento Geral sobre a Proteção de Dados[93]*.*

Ao seu turno, a Rede Iberoamericana de Proteção de Dados, igualmente, iniciou em maio de 2023 uma ação coordenada para garantir a proteção de direitos e liberdades dos indivíduos afetados pelo ChatGPT e, consequentemente, por sistemas de IA generativa que utilizam LLMs. Nas exatas palavras da autoridade iberoamericana[94]:

> La RIPD considera que este servicio, que brinda respuestas directas a preguntas formuladas por los usuarios, desarrollado por la empresa Open AI, L.L.C., puede conllevar riesgos para los derechos y libertades de los usuarios en relación con el tratamiento de sus datos personales, los que abarcan aspectos tales como, los fundamentos legales para dichos tratamientos, la información que sobre los tratamientos se brinda al usuario, el ejercicio de los derechos reconocidos en las normativas de protección de datos, las posibles transferencias de datos personales a terceros sin contar con el consentimiento de los titulares, el no contar con medidas de control de edad para impedir que menores accedan a su tecnología así como no saber si cuenta con adecuadas medidas de seguridad para la protección y confidencialidad de los datos personales recabados (RIPD, 2023).

[93] Conf. www.aepd.es/prensa-y-comunicacion/notas-de-prensa/aepd-inicia-de-oficio-actuaciones-de-investigacion-a-openai. Acesso em 23 de janeiro de 2024.
[94] Conf. www.redipd.org/es/noticias/autoridades-red-iberoamericana-de-proteccion-de-datos-personales-inician-accion-chatgpt. Acesso em 23 de janeiro de 2024.

Portanto, ficou demonstrando que a IA generativa faz parte de uma estrutura de inteligência artificial, que se utiliza, notadamente, de *deep learning* e, mais precisamente, de *LLMs*, para processar dados rotulados e não-rotulados, a fim de haja uma criação de respostas a partir do aprendizado destes dados. Todavia, há evidente preocupação com os dados utilizados para treinamento destes LLMs, saltando da lógica *garbage in, garbage out*, para uma lógica de minimização de vieses, de maneira que exista mais precisão e qualidade nos resultados obtidos, combatendo os riscos de qualidade, discriminação e inovação, elencados por HACKER (2021).

4.3) Proteção de dados e IA Generativa: proteção de dados e IA Generativa:

PHILIPP HACKER, acertadamente, ensina que os dados de treinamento ainda representam um lugar inóspito no mundo jurídico e sua importância central para técnicas de aprendizado de máquina, no entanto, sugere que, ao contrário de uma visão generalizada, não se trata tanto de uma regulamentação de algoritmos, mas de uma regulamentação de proteção de dados necessários notadamente os dados de treinamento, para enfrentar os riscos centrais, quais sejam, os i) riscos para a qualidade dos dados; ii) riscos de discriminação; e iii) riscos de inovação (Hacker, 2021).

Segundo a autoridade nacional de dados italiana (Garante della Privacy) a privacidade e a proteção de dados representam obstáculos legais críticos ao desenvolvimento e implantação de IA generativa, conforme exemplificado pela proibição temporária do ChatGPT, corroborando o risco da inovação através do bloqueio lastreado na proteção de dados pessoais.

Num nível abstrato, um LLMs preserva a privacidade se divulgar informações confidenciais em contextos apropriados e apenas a indivíduos autorizados. A privacidade e a proteção de dados não são variáveis binárias e, portanto, qual é o contexto certo ou os destinatários certos da informação é uma questão de debate. No contexto dos LLMs, esses debates são ainda mais complicados devido aos diversos propósitos, aplicações e ambientes em que operam (Hacker *et al.* 2024)

Os LLMs estão expostos a violações de privacidade e proteção de dados devido ao treinamento generalizado em dados (parcialmente) pessoais, à memorização de dados de treinamento e a ataques de inversão (Hacker et al. 2024). A memorização de dados pode ocorrer através do ajuste excessivo de parâmetros abundantes a pequenos conjuntos de dados, o que reduz a capacidade de generalização para novos dados, ou através da generalização otimizada de distribuições de dados de cauda longa. Quando os dados de treinamento memorizados contêm informações pessoais, os LLMs podem vazar dados e divulgá-los diretamente.

Quando os dados de treinamento não são memorizados, as informações pessoais ainda podem ser inferidas ou reconstruídas por atores mal-intencionados usando ataques de inversão de modelo, que fazem engenharia reversa dos dados de entrada para revelar informações privadas, ferindo a privacidade dos utilizadores (Hacker *et al.,* 2024 *apud* Fredrikson *et al.,* 2015).

Contra isto, as estratégias existentes de preservação da privacidade, tais como a “higienização de dados” e a “privacidade diferencial”, proporcionam uma proteção limitada da privacidade quando aplicadas aos LLMs. Isto levanta a questão de saber se, e como, os dados pessoais podem ser processados para formar LLMs – uma questão particularmente espinhosa no que diz respeito a dados sensíveis. Além disso, os usuários podem inserir informações privadas por meio de prompts, que podem reaparecer em outros casos. De mais, alguns utilizadores serão menores, aos quais se aplicam regras específicas de proteção de dados.

Assim, seguimos ancorados na doutrina de PHILIPP HACKER quanto aos problemas específicos na interseção de IA generativa e proteção de dados, quais sejam, para HACKER *et al* (2024) propagam-se sete problemas principais na interseção da proteção de dados e sistemas de IA com base em LLMs: i) a base jurídica apropriada para a treinamento de sistemas de IA com dados pessoais; ii) a base jurídica apropriada para o processamento de solicitações; iii) requisitos de informação; iv) inversão de modelos, vazamento de dados e direito ao apagamento; v) tomada de decisão automatizada; vi) proteção de menores; e vii) limitação de finalidade e minimização de dados

Para a formação de uma base jurídica apropriada para a treinamento de sistemas de IA com dados pessoais, segundo HACKER et al. (2024) fundamental que se observe dois pilares, quais seja, consentimento e proporcionalidade.

Contudo, frise-se que todas as operações de tratamento de dados pessoais – seja armazenamento, transferência, treinamento – necessitam de uma base jurídica nos termos do artigo 6.º do RGPD e artigo 7.º da LGPD, como mediante o fornecimento de consentimento pelo titular ou quando necessário para atender aos interesses legítimos do controlador ou de terceiros, exceto no caso de prevalecerem direitos e liberdades fundamentais do titular que exijam a proteção dos dados pessoais.

Isso se aplica, inclusive, para empresas sem estabelecimento na EU ou no território brasileiro, o que abrange a conformidade das grandes *big techs* ao quadro geral de proteção de dados que, geralmente tem sua sede nos Estados Unidos da América, China ou Índia, à teor do artigo 3.º da LGPD e do RGDP.

Como dito, a base jurídica que garante a licitude do tratamento de dados mais proeminente do RGPD e na LGPD é o consentimento. No entanto, para *big data* e, consequentemente LLMs, incluindo informações pessoais de um vasto grupo de pessoas previamente desconhecidas pelos desenvolvedores, obter o consentimento válido de cada indivíduo geralmente não é uma opção devido aos custos de transação proibitivos.

Além disso, é difícil conciliar o uso de LLMs com conjuntos de dados extraídos da Web e aplicações imprevisíveis com consentimento informado e específico, para HACKER et al. (2024). Ao mesmo tempo, exigir que os titulares dos dados sejam informados sobre a utilização dos seus dados pessoais pode retardar o desenvolvimento de LLMs. Assim, por razões jurídicas e econômicas, a formação em IA pode normalmente basear-se no teste de proporcionalidade, previsto do Artigo 6(1)(f) do RGPD e artigo 7º, IX da LGPD, segundo o qual os interesses legítimos do responsável pelo tratamento (ou seja, a entidade de desenvolvimento) justificam o processamento, a menos que sejam anulados

pelos direitos e liberdades dos titulares dos dados (ou seja, as pessoas cujos dados são utilizados).

Particularmente em relação ao princípio da proporcionalidade, ensina-nos o Prof. Doutor REIS NOVAIS que a o princípio em voga deve verificar os sacrifícios e benefícios quando do choque entre dois direitos:

> Neste controle de proporcionalidade, aquilo que se avalia, que se compara ou que se põe em relação, são os sacrifícios (custos) impostos ao direito fundamental contraposto aos benefícios (vantagens) produzidos na obtenção do fim visando com a restrição (Novais, 2019, p. 250).

Em idêntico sentido, o Professor Doutor GILMAR MENDES destaca que, para se efetuar restrições aos direitos fundamentais, é pressuposto lógico identificar o âmbito de proteção do direito:

> [...] o exame das restrições aos direitos fundamentais pressupõe a identificação do âmbito de proteção do direito, vez que esse processo não pode ser fixado em regras gerais, exigindo, para cada direito específico, determinado procedimento (Mendes, 2017, p. 174).

Portanto, a aplicação do princípio da proporcionalidade deverá utilizar-se da ponderação entre os sacrifícios e as vantagens de determinado modelo de IA, especialmente no choque do direito à licitude do tratamento com base no consentimento em contraposição à liberdade de utilização de dados pelos desenvolvedores lastreado em interesses igualmente legítimos.

Geralmente, aplicativos benéficos socialmente falarão a favor dos desenvolvedores; do mesmo modo, é pouco provável que o controle prevaleça se a utilização dos dados para fins de formação em IA puder ser razoavelmente esperada pelos titulares dos dados, pelo que este último critério raramente será cumprido. Por outro lado, a natureza e o âmbito do tratamento, o tipo de dados (sensíveis ou não), o grau de transparência e controlo dos titulares dos dados e outros fatores podem fazer pender a ponderação no teste de proporcionalidade noutra direção (Hacker *et al.*, 2024).

Assim, para modelos de IA estreitamente adaptados, baseados em estratégias de aprendizagem supervisionada, pode-se argumentar que a formação em IA não é particularmente útil, uma vez que, geralmente, não revela

quaisquer informações novas sobre os próprios titulares dos dados (Hacker 2023c *apud* Kirrane 2019). No entanto, esta posição é difícil de manter em relação aos LLMs pois estes modelos são geralmente utilizados por milhões de intervenientes diferentes, e foi demonstrado que os modelos também revelam dados pessoais através de fuga de dados, representando um desafio ainda maior em cenários de ajuste fino (Hacker *et al*., 2024 *apud* Nicolas Carlini *et al*., 2023).

Outrossim, a utilização de dados sensíveis por LLMs, essenciais em diversas áreas de desenvolvimento de IA generativa como na medicina e produção de medicamentos, tem um contorno especial, na medida em que o consentimento é a regra inafastável para o tratamento destes, como se depreende do artigo n.º 11 da LGDP e do 9.º da RGPD, não havendo uma cláusula de ponderação, como no caso de dados não-sensíveis, como dito adrede.

Para o saudoso professor DANILO DONEDA, uma característica intrínseca da sensibilidade dos dados, como merecedora de uma tutela mais robusta, é a possibilidade potencial de utilização discriminatória e, portanto, violadora de direitos fundamentais. Senão vejamos:

> O regime adotado em relação aos dados sensíveis varia de acordo com as concepções a este respeito em cada ordenamento. Na verdade, deve-se ter em conta que a diferenciação conceitual dos dados sensíveis atende à uma necessidade de estabelecer uma área na qual a probabilidade de utilização discriminatória da informação é potencialmente maior [...]. (Doneda, 2019, p. 144)

No âmbito da UE, a decisão do TJUE no caso Meta v. Bundeskartellamt[95], firmou a jurisprudência no sentido de que as informações não precisam referir-se diretamente a categorias protegidas – como origem étnica ou racial, religião, idade ou saúde – para serem abrangidas pelo artigo 9.º da RGPD.

O Tribunal considerou que não importa, por exemplo, se a pessoa traçada é um utilizador do Facebook ou não, pelo contrário, do ponto de vista da

[95] Conf. Tribunal de Justiça da União Europeia. Decisão n.º CJEU, C-252/21, *Meta vs. Bundeskartellamt*, ECLI:EU:C:2023:537. 2021.

RGPD, o que é decisivo é a capacidade do responsável pelo tratamento de inferir características sensíveis com base nos dados disponíveis, independentemente de o operador pretender fazer essa inferência. Logo, este entendimento mais amplo lança uma ampla rede para a aplicabilidade do artigo 9 do RGPD, à medida que as técnicas de aprendizagem automática permitem cada vez mais a dedução de categorias protegidas a partir de pontos de dados que de outra forma seriam inócuos.

Assim, em muitos casos relativos a formatos de grandes volumes de dados, a possibilidade hipotética de inferir dados sensíveis coloca potencialmente o tratamento, por exemplo, para fins de formação em IA, no âmbito do artigo 9.º. (2) RGPD, ou seja, fora do consentimento explícito, tal exceção, no entanto, muitas vezes não estará disponível.

Todavia, no diálogo de fontes entre a proteção de dados e a AIA poderá existir uma solução intermédia., ou seja, está previsto no Artigo 10(5) AIA, para equilibrar o interesse da sociedade na formação e desenvolvimento de IA socialmente benéfico com a proteção dos direitos e liberdades individuais, especialmente em áreas cruciais como medicina, educação ou emprego. Embora a exceção TDM preveja um quadro específico para a utilização de material protegido por direitos de autor para fins de formação em IA, tais regras são, infelizmente, totalmente inexistentes no âmbito do RGPD.

No que concerne à base jurídica para solicitações contendo dados pessoais, HACKER *et al.* (2024) adverte que temos que distinguir fundamentalmente duas situações. Primeiro, os usuários podem incluir informações pessoais sobre si mesmos em *prompts*, por exemplo, quando solicitam a um LLM que redija um e-mail sobre um evento, compromisso ou tarefa específico. Neste caso, o consentimento pode de fato funcionar como base legal, uma vez que os utilizadores têm de registar-se individualmente para o produto LLM, sendo que durante esse procedimento, os responsáveis pelo tratamento podem consentir, respeitando as condições de consentimento válido nos termos do artigo 4.º, n.º 11, e do artigo 7.º do RGPD e artigos n.º 7, I e 11, I da LGPD.

O segundo cenário, conforme HACKER *et al.* (2024) diz respeito a solicitações que contêm informações pessoais sobre terceiros, ou seja, não sobre a pessoa que digita o *prompt*, haja vista que eles podem incluir inadvertidamente detalhes pessoais de outras pessoas se a tarefa em questão envolver esses terceiros, e esperam que o modelo de linguagem forneça respostas personalizadas. Todavia, os utilizadores não podem consentir validamente para outra pessoa (a menos que tenham sido explicitamente mandatados por essa pessoa para fazer exatamente isso, o que é improvável).

Consequentemente, um problema semelhante ressurge como no cenário de treinamento ou ajuste fino de IA, com a diferença adicional de que a informação é fornecida e o processamento é iniciado pelo usuário, e não pelos desenvolvedores. Embora o usuário possa ser considerado o único controlador ou controlador conjunto junto com a empresa que opera o LLM (Artigo 4 (7) do GDPR), para o armazenamento inicial e transferência do aviso (ou seja, redação e envio do aviso), qualquer memorização adicional ou vazamento de dados está sob o controle exclusivo da entidade que opera o LLM e será provavelmente considerado o único responsável pelo tratamento e, portanto, a parte responsável, segundo o artigo 5.º, n.º 2, do RGPD.

No que tange às questões de informação, os principais obstáculos para LLMs compatíveis com a proteção da dados são os artigos 12º a 15º do RGPD, que detalham as obrigações relativas às informações que devem ser fornecidas aos titulares dos dados. Estes artigos representam um desafio único para os LLMs devido à natureza e ao âmbito dos dados que processam (Hacker *et al.,* 2024).

Ao considerar os dados recolhidos na Internet para fins de formação, a aplicabilidade do artigo 14.º do RGPD é crucial, tendo em conta que este comando aborda a necessidade de transparência nos casos em que os dados pessoais não são recolhidos diretamente dos indivíduos em causa, tais como as finalidades do tratamento a que os dados pessoais se destinam, bem como o fundamento jurídico para o tratamento; as categorias dos dados pessoais em questão e os destinatários ou categorias de destinatários dos dados pessoais, se os houver.

No entanto, a viabilidade de informar individualmente aqueles cujos dados fazem parte do conjunto de formação é muitas vezes impraticável devido ao grande esforço necessário, potencialmente isentando-o ao abrigo do artigo 14.º, n.º 5, alínea b), do RGPD[96]. Por outro lado, o tratamento de dados pessoais submetidos pelos próprios utilizadores numa interface de chat (prompts) não está sujeito a tais isenções. O artigo 13.º do RGPD exige explicitamente que os titulares dos dados sejam informados de vários aspectos fundamentais, incluindo as finalidades do tratamento, a base jurídica do tratamento e quaisquer interesses legítimos prosseguidos.

HACKER *et al.* (2024) ressalta que o equilíbrio entre os desafios práticos de conformidade e os direitos dos titulares dos dados é delicado, vez que, embora o conceito de esforço desproporcional nos termos do artigo 14.º, n.º 5, do RGPD apresente uma isenção potencial, continua a ser um ponto controverso, especialmente no que diz respeito à recolha e processamento de dados para fins comerciais. A este respeito, o responsável pelo tratamento de dados, conforme definido no artigo 4.º, n.º 7, do RGPD, deve documentar meticulosamente as considerações feitas ao abrigo desta disposição. Esta documentação é um aspecto crucial do princípio de responsabilização consagrado no artigo 5.º, n.º 2, do RGPD. Além disso, na nossa opinião, os documentos relativos aos métodos de recolha de dados de formação devem ser tornados acessíveis ao público, reforçando o compromisso com os princípios do RGPD.

Analogamente, no âmbito brasileiro, a LGPD através do artigo n.º18 dispõe que o titular de dados pessoais o direito de obter do controlador a qualquer momento informações sobre o tratamento dos dados, bem como a retirada do consentimento, opondo-se ao tratamento com fundamento em uma das hipóteses de dispensa de tratamento, em caso de descumprimento da

[96] Art. 14.º, n.º 5, alínea b) - Se comprove a impossibilidade de disponibilizar a informação, ou que o esforço envolvido seja desproporcionado, nomeadamente para o tratamento para fins de arquivo de interesse público, para fins de investigação científica ou histórica ou para fins estatísticos, sob reserva das condições e garantias previstas no artigo 89.o, n.o 1, e na medida em que a obrigação referida no n.o 1 do presente artigo seja suscetível de tornar impossível ou prejudicar gravemente a obtenção dos objetivos desse tratamento. Nesses casos, o responsável pelo tratamento toma as medidas adequadas para defender os direitos, liberdades e interesses legítimos do titular dos dados, inclusive através da divulgação da informação ao público;

LGPD. Desta maneira, uma vez não obedecido os princípios elencados no artigo 6º, como *verbi gratia*, o da finalidade, da transparência, não discriminação poderá o titular opor-se ao uso de seus dados em sistemas de IA e LLMs.

Noutro giro, a conformidade LLMs com a RGPD e a LGPD fica ainda mais complicada com preocupações sobre a reconstrução de dados de treinamento do modelo e vazamentos de dados não intencionais, especialmente à luz do direito ao esquecimento (ou direito ao apagamento) nos termos do artigo 17 (Hacker *et al*. 2024)

HACKER *et al.* (2024) ensina que os ataques de inversão ou reconstrução de dados de treinamento referem-se a técnicas pelas quais, por meio de ataques específicos, os dados dos indivíduos utilizados no treinamento desses modelos podem ser extraídos ou inferidos. Da mesma forma, o problema de memorização, que faz com que os LLMs produzam potencialmente dados pessoais contidos nos dados de treinamento, pode ser invocado para qualificar os próprios LLMs como dados pessoais.

Se um LLM utilizar, portanto, dados pessoais, isso implica que os titulares dos dados poderiam, em tese, invocar o seu direito ao apagamento nos termos do artigo 17.º do RGPD. Este direito, também conhecido como “direito a desindexação”, permite que os indivíduos solicitem a eliminação dos seus dados pessoais em condições específicas. No contexto dos LLMs, isto poderia levar a exigências sem precedentes para a eliminação do próprio modelo, caso se estabelecesse que o modelo contém ou constitui dados pessoais dos indivíduos.

Para HACKER *et al.* (2024) tal cenário apresenta desafios significativos para o campo da IA e do aprendizado de máquina. A praticidade de atender a um pedido de apagamento neste contexto está repleta de complexidades técnicas e jurídicas. A exclusão de um modelo, especialmente um que tenha sido amplamente distribuído ou implantado, pode ser um desafio tecnológico e ter implicações significativas na utilidade e funcionalidade do sistema. Além disso, esta abordagem levanta questões sobre o equilíbrio entre os direitos individuais e os benefícios mais amplos das tecnologias de IA.

A eliminação de modelos inteiros, com um potencial de necessidade econômica subsequente de requalificar todo o modelo, também levanta questões

complexas relativas à sustentabilidade ambiental, dado o enorme consumo de energia e água da (re)formação de LLMs.

Embora os produtores de LLM, como a OpenAI, aleguem cumprir o direito ao apagamento, não está claro como o podem fazer porque as informações pessoais podem ser transmitidas de múltiplas formas num LLM, o que aumenta a complexidade da identificação e isolamento de pontos de dados específicos, especialmente quando os dados não são apresentados de forma formato estruturado (por exemplo, números de telefone).

Os dados incorporados durante a fase de treinamento podem permear os resultados gerados por determinados modelos de aprendizado de máquina, criando um cenário onde os dados de treinamento originais, ou informações vinculadas aos dados eliminados, podem ser inferidos ou "vazados", minando assim a integridade do processo de exclusão e perpetuando potenciais violações de privacidade (Hacker *et al*., 2024 *apud* De Cristofaro, 2020). No mínimo, isto aponta para a necessidade de estratégias mais robustas e abrangentes para abordar a privacidade de dados na área operacional dos LLMs, que poderá se valer do princípio do *privacy by design*.

Outra vital preocupação entre proteção de dados e sistemas de IA diz respeito às decisões automatizadas, já direcionada em capitulo específico deste estudo. Entretanto, particularmente em relação os LLMs, vale ressaltar que nos casos em que os LLMs são utilizados para avaliação, como no recrutamento ou na pontuação de crédito, a importância deste regulamento torna-se ainda mais significativa.

Visando tutelar, portanto, a proteção de dados pessoais utilizados nestes modelos de treinamento, para garantir a qualidade através de treino, validação e teste, o AIA, através do artigo 10.º, traz uma série de regras para o treinamento de dados de formação. De tal sorte que os sistemas de IA de risco elevado que utilizem técnicas que envolvam o treino de modelos com dados devem ser desenvolvidos com base em conjuntos de dados de treino, validação e teste que cumpram os critérios de qualidade referidos nos números 2 a 5.

Ou seja, os conjuntos de dados de treino, validação e teste devem estar sujeitos a práticas adequadas de governação e gestão de dados, sendo que tais práticas dizem nomeadamente respeito: a) às escolhas de conceção tomadas; b) à recolha de dados; c) às operações de preparação e tratamento de dados necessárias, tais como anotação, rotulagem, limpeza, enriquecimento e agregação; d) à formulação dos pressupostos aplicáveis, nomeadamente no que diz respeito às informações que os dados devem medir e representar; e) à avaliação prévia da disponibilidade, quantidade e adequação dos conjuntos de dados que são necessários; f) ao exame para detectar eventuais enviesamentos; g) à identificação de eventuais lacunas ou deficiências de dados e de possíveis soluções para as mesmas[97].

Ademais, os conjuntos de dados de treino, validação e teste devem ser pertinentes, representativos, isentos de erros e completos, bem como devem ter as propriedades estatísticas adequadas, nomeadamente, quando aplicável, no tocante às pessoas ou grupos de pessoas em que o sistema de IA de risco elevado se destina a ser utilizado. Os conjuntos de dados de treino, validação e teste devem ter em conta, na medida do necessário para a finalidade prevista, as características ou os elementos que são idiossincráticos do enquadramento geográfico, comportamental ou funcional específico no qual o sistema de IA de risco elevado se destina a ser utilizado.

Outrossim, consoante item 5) do artigo n.10º, na medida do estritamente necessário para assegurar o controlo, a deteção e a correção de enviesamentos em relação a sistemas de IA de risco elevado, os fornecedores desses sistemas podem tratar categorias especiais de dados pessoais a que se refere o artigo 9.º, n.º 1, do Regulamento (UE) 2016/679, o artigo 10.º da Diretiva (UE) 2016/680 e o artigo 10.º, n.º 1, do Regulamento (UE) 2018/1725, assegurando salvaguardas adequadas dos direitos fundamentais e liberdades das pessoas singulares.

Destarte, isso inclui limitações técnicas à reutilização e utilizar medidas de segurança e preservação da privacidade de última geração, tais

[97] Conf. art. 10 do Regulamento para Inteligência Artificial da União Europeia.

como a pseudonimização ou a cifragem nos casos em que a anonimização possa afetar significativamente a finalidade preconizada.

No Brasil, o PL n.º2338/23 (PLIA) traz previsão mais tímida sobre dados de treinamento dizendo que, como parte do *governance* dos sistemas de IA, os agentes de IA estabeleceram estruturas de governação e processos internos aptos a garantir a segurança dos sistemas e o atendimento dos direitos das pessoas afetadas, que incluíram, ao menos, a adoção de parâmetros adequados de separação e organização de dados para treinamento, teste e validação dos resultados dos sistemas, deixando subjetivamente à escolha das empresas o estabelecimento destes parâmetros, em visível adoção de um modelo autorregulatório.

Destarte, pertinente apontar que, em relação à prevenção de discriminações, o supracitado comando deverá ser lido em conjunto com o que determina o inciso IV do art. 20 do PLIA, que impõe, para além da medida acima, novas medidas de gestão para mitigar e prevenir vieses discriminatórios, que incluem a avaliação dos dados, para controle de vieses cognitivos humanos que possam afetar a coleta e organização dos dados, bem como medidas corretivas para evitar incorporação de vieses sociais estruturais.

De mais, a composição de equipe responsável pela concepção e desenvolvimento do sistema deverá ser inclusiva, orientada pela busca da diversidade, a fim de minimizar riscos de qualidade e discriminação inerentes à IA generativa.

Uma ilustração pertinente é fornecida pela recente decisão do TJUE no caso SCHUFA[98], onde o Tribunal determinou que a geração automatizada de um valor de probabilidade relativo à capacidade futura de um indivíduo de pagar compromissos por uma agência de informações de crédito constitui uma “tomada de decisão individual automatizada”. ', conforme definido no artigo 22.º do RGDP. Segundo o Tribunal, isto pressupõe, no entanto, que este valor de probabilidade

[98] Conf. Tribunal de Justiça da União Europeia. CJEU, C-634/21, QG vs. SCHUFA, ECLI:EU:C:2023:957, para. 73. 2021.

influencia significativamente a decisão de um terceiro de celebrar, executar ou terminar uma relação contratual com esse indivíduo.

Extrapolando esta decisão, a avaliação ou classificação automatizada de indivíduos pelos LLMs constituirá uma tomada de decisão automatizada se for de suma importância para a decisão em questão – mesmo que um ser humano a aprove posteriormente. As implicações jurídicas disso são profundas. As isenções à proibição geral de tal tomada de decisão automatizada estão limitadas a cenários em que exista uma lei específica que permita o processo, consentimento explícito, ou onde o tratamento automatizado seja necessário para fins contratuais, nos termos do artigo 22.º, n.º 2, do RGPD.

Entrementes, a implantação de LLMs suscitaram preocupações significativas relativamente ao conteúdo adequado à idade, especialmente tendo em conta o potencial de geração de resultados que podem não ser adequados para menores. Nos termos do artigo 8.º, n.º 2, do RGPD, o responsável pelo tratamento deve realizar esforços razoáveis para verificar que o consentimento é dado ou autorizado pelo titular da responsabilidade parental sobre a criança, tendo em consideração a tecnologia disponível.

A LGPD, por sua vez, dedica a Seção III da lei inteira para proteção de dados de crianças e adolescentes. De forma análoga ao RGPD, a lei brasileira diz, entre outros, que o tratamento de dados pessoais de crianças deverá ser realizado com o consentimento específico e em destaque dado por, pelo menos, um dos pais ou pelo responsável legal e o controlador deve realizar todos os esforços razoáveis para verificar que o consentimento foi dado pelo responsável pela criança, consideradas as tecnologias disponíveis.

Um exemplo notável de intervenção regulatória neste contexto é a ação tomada pela Autoridade Italiana de Proteção de Dados (Garante per la Protezione dei Dati Personali –- GPDP). Em 30 de março de 2023, o GPDP impôs uma restrição temporária ao processamento de dados de usuários italianos pela OpenAI, com ênfase particular na proteção de menores. Esta medida ressalta o crescente escrutínio por parte das autoridades de proteção de dados sobre as implicações dos LLMs no contexto da proteção grupos vulneráveis, especialmente crianças (Hacker *et al.,* 2024 *apud* Malgieri, 2023).

Em resposta a estas preocupações, a OpenAI, por exemplo, implementou medidas destinadas a melhorar a proteção dos menores. Estas incluem o estabelecimento de uma diferença de idade e a integração de ferramentas de verificação de idade. A eficácia e a robustez destas ferramentas, no entanto, continuam a ser uma área de grande interesse e de avaliação contínua, especialmente no panorama em rápida evolução da IA e da proteção de dados.

Finalizando, em relação à limitação de finalidades e minimização de dados, os responsáveis pelo tratamento de dados devem recolher dados pessoais apenas se forem relevantes e necessários para uma finalidade específica, *ex vi* do artigo 5.º, alíneas b) a c). Já o AIA reflete isto, exigindo uma avaliação da quantidade e adequação dos dados, conforme o artigo 10.º, alínea e). A LGPD, igualmente, prevê como princípio elencado no rol do artigo 6º que a necessidade implica a limitação do tratamento ao mínimo necessário para realização das finalidades.

No entanto, para HACKER (2023) limitar a gama indefinida de finalidades dos LLMs, que necessitam de dados extensos para uma formação eficaz, pode ser inútil e contraproducente. Uma abordagem para lidar com a calibração de dados para aplicativos LLM abertos é exigir que os desenvolvedores treinem modelos em conjuntos de dados menores e aproveitem habilidades de aprendizagem de poucos ou nenhum disparo.

Contudo, como alternativa à imposição de restrições ao conjunto de dados, poderia ser mais benéfico reforçar as medidas de preservação da privacidade proporcionalmente ao tamanho do conjunto de dados. Por exemplo, em vez de depender apenas da pseudo-anonimização e da encriptação (artigo 10.º da AIA), os fornecedores de LLM devem implementar métodos como a privacidade diferencial para combater ataques adversários a grandes conjuntos de dados (Hacker *et al. 2024 apud* Shi *et al.,* 2022)

Desta maneira, falar em inteligência artificial generativa é ter em conta o imenso desafio à proteção de dados pessoais perante os LLM's, especialmente os riscos de qualidade, de discriminação e de inovação, por isso enfrentamento destes riscos por um sistema regulatório deverá consagrar uma base jurídica

que dê legitimidade ao tratamento de dados e solicitações de usuários; assegure o direito à informação e ao apagamento e retificação de dados; que proteja os menores de idade; que limite a finalidade e cumpra a minimização de dados pessoais e; por fim, que assegure salvaguardas às decisões automatizadas, com imposição de responsabilidade.

O constitucionalismo digital aqui inspirou este quadro-legal suprarreferido, na medida em que há um estabelecimento dos limites legais da legislação e a expectativa de devido cumprimento por todos os atores estatais e não-estatais envolvidos, nomeadamente através do *enforcement* das autoridades de controle, contemplado uma ideia de regulação tradicional (dimensão interna ou doméstica); somado uma dimensão internacional e comunitária (como no caso da UE), onde uniformizam-se regras aplicadas a diversos países; e, por fim, à senda privada ou não-estatal, pautada ora na autorregulação, ora na autorregulação regulada (Celeste, 2019).

5) Modelos regulatórios Brasil e União Europeia:

5.1) Regulação: constitucionalismo digital como parâmetro de regulação:

A nova esfera pública digital, portanto, para além dos efeitos benéficos do avanço tecnológico, está eivada de sérios problemas, como propagação intensa de conteúdos ilegais, como as notícias falsas e desinformação, bem como as perfilização de usuários, vigilância e o uso massivo de dados pessoais para diversos propósitos, especialmente comerciais.

De mais, o avanço desregulado na utilização da inteligência artificial está no centro do debate, especialmente das chamadas IAs generativas, chegando ao extremo de "personalidades tech", como Elon Musk, dizer que é preciso frear e impor limites severos à inteligência artificial, sob pena de não retorno ao *status quo*, havendo um potencial de danos da magnitude de armas nucleares[99].

[99] "AI more dangerous than nukes". Conferir https://www.reuters.com/video/watch/idRCV004NTE; acesso em 10 de julho de 2023.

Assim, soluções e endereçamentos dos problemas típicos do ciberespaço devem ter o condão de ultrapassar velhas noções de territorialidade, soberania e morosidade legislativa, isto pois a extraterritorialidade das empresas e usuários é ínsita ao modelo digital, diminuindo, por consequência, a ideia de soberania e submissão à lei determinado Estado. Em paralelo, a velocidade em que as novas tecnologias são criadas e passam a ser utilizadas na vida diária dos usuários é demasiadamente superior à normatividade estatal típica.

Por isso, a normatividade digital e as relações entre os sistemas tecnológico e digital e o sistema jurídico configuram, atualmente, verdadeiros paradoxos para o Direito Público, em especial, para o Direito Constitucional. Nos últimos anos, assistimos a uma implacável desconstrução da ordem jurídico-constitucional, perpetrada pelos factos carreados para o sistema normativo pela Internet e Novas Tecnologias (Castro, 2023).

Para professora doutora RAQUEL BRÍZIDA CASTRO, em linha com o defendido por SOUSA (2022) e DE GREGORIO (2022), a regulação do espaço digital deve fundar-se num inelutável *Pluralismo Normativo Multinível – Plurinormativismo Tecnológico e Digital*-, dissonante dos pilares tradicionais do Direito Constitucional, mas particularmente atrativo no ciberespaço (Castro, 2023, p.53).

SOUSA (2022) , na mesma linha de raciocínio, afirma que a regulação das plataformas digitais é matéria de Direito Público porque possui significação política e importam à coletividade no seu todo e não apenas a um determinado e restrito grupo de pessoas.

O renomado Professor alemão GUNTHER TEUBNER realça a crise do constitucionalismo moderno que tem peculiar dificuldade em compreender e dar respostas adequadas à “nova questão constitucional”, tais como aquelas já apontadas nessa investigação, perpetradas no âmbito da esfera pública digital, como a interferência massivas na esfera privada decorrente da coleta e retenção de dados por organizações privadas:

> Violações de direitos humanos por empresas multinacionais; decisões controversas da OMC que, em nome do livre comércio global, ameaçam a proteção ao meio ambiente e à saúde; doping esportivo; corrupção na medicina e na ciência; ameaças

> à liberdade de expressão por intermediários privados na internet; interferências massivas na esfera privada decorrente da coleta e retenção de dados por organizações privadas [...] (Teubner, 2020, p. 41).

Isso se deve, em razão de alguns fatores, segundo TEUBNER (2020), como a dinâmica incontrolável dos mercados de capitais globais, o evidente poder de empresas transnacionais, especialmente as *big tech*, e a dominância de *experts* não legitimados em extensas *epistemic communities* não informadas pelo direito, que conduzem, assim, para um constitucionalismo transnacional.

Vejamos:

> Nesse caso, três fenômenos situam-se no primeiro plano (1) a desconstitucionalização do Estado Nacional é desencadeada pelo deslocamento de funções de governo para o âmbito transnacional, bem como pela assunção de parte dessas funções por atores não estatais; (2) efeitos extraterritoriais da atuação dos Estados Nacionais permitem o surgimento de um Direito que carece de legitimação democrática; e, por fim, (3) a inexistência de mandato democrático para a *governance* transnacional (Teubner, 2020, p. 49).

SOUSA (2022) *apud* INGOLF PERNICE, corrobora a visão de TEUBNER (2017), na medida em que reconhece que o constitucionalismo digital é uma nítida manifestação inspirada no constitucionalismo multinível que, por sua vez, trata-se de uma teoria normativa baseada num modelo organizativo para instituições transnacionais partindo de fontes locais e transnacionais que se harmonizam e interpenetram, para o exercício de poder em camadas transparentes e obedientes a um processo organizado e igualitário.

> Se isto é verdade, não é menos verdade que o constitucionalismo digital promove a humanização do mundo digital e, por isso, ao complementarem-se assistimos a um processo de assimilação que promove o surgimento do constitucionalismo digital multinível (Sousa, 2022, p.106).

Destarte, o igualmente laureado Professor WOLFGANG HOFFMANN-RIEM aduz que a transformação digital está encontrando um novo ajuste na relação entre direito privado e o direito público, especialmente como resultado de medidas anteriores de desregulamentação e privatização, pelo que parece sustentável a transferência do poder de agir para instituições privadas (Hoffmann-Riem, 2021).

Este movimento vai ao encontro de respostas para àquelas questões essenciais postas inicialmente sobre i) soberania e ii) território, em razão da extraterritorialidade da presença e influência destes novos atores privados da tecnologia, que devem passar a ser agentes atuantes, defensores e promotores dos direitos fundamentais dentro de seu espectro de atuação, como um imperativo do seu papel público na formação da comunidade digital, deixando a noção de partes meramente privadas, atreladas ao consentimento e ao *pacta sund servanda* dos termos e condições de adesão aos serviços e produtos ofertados pela plataforma, para trás.

Contudo, HOFFMANN-RIEM (2021), com precisão, ensina que a autorregulação pelos *players* privados deve se ater às balizas legais e constitucionais, pelo que autodeterminação privada e da autorregulação não altera, portanto, a tarefa do Estado como garante em assumir uma sua responsabilidade pela salvaguarda do bem individual e comum, advogando pela utilização do modelo de autorregulação ou regulação regulada:

> Os particulares – protegidos pelas liberdades civis- são, em princípio livres para perseguir seus interesses e especificar seus cálculos de benefício. No entanto, não estão completamente isentos de consideração pelos interesses dos outros e pelo bem comum. Se necessário, a lei pode ou deve estabelecer uma estrutura para garantir o exercício socialmente aceitável da liberdade. A grande importância da autodeterminação privada e da autorregulação não altera, portanto, a tarefa do Estado como "Estado Garantidor" de assumir uma "responsabilidade garantidora" pela salvaguarda do bem individual e comum também por lei (Hoffmann-Riem, 2021, p. 135-136).

Na mesma linha de raciocínio, o pesquisador SIMÃO SOUSA, como já exposado, advoga que o esforço para a constitucionalização das plataformas digitais deve ser tomado em direção a uma governação digital multinível, com atuação todos os níveis, seja internacional e comunitário, seja doméstico, mas fundamentalmente na autorregulação privada, objetivando também o estabelecimento de mecanismos e princípios comuns, compatíveis e atualizados ao espaço digital, ou seja, uma mínima e aplicável uniformidade de regras (Sousa, 2022).

Desta forma, autorregulação ou regulação regulada seria aquela oriunda tanto das autoridades públicas, quanto de instituições privadas. Aquelas contam com os serviços de regulação prestados pelos membros da sociedade em relativa autonomia – desde que observado parâmetros constitucionais e legais-, para a solução de problemas oriundos do mundo digital, em especial sua capacidade de enfrentamento de respostas em velocidade superior ao Estado e a capacidade técnica dos envolvidos em prover respostas mais acuradas.

Esta abordagem parece-nos correta ao considerar o papel público das instituições privadas de tecnologia do ciberespaço e sua influência significativa no debate público e na democracia de forma geral. Nesta linha raciocínio é também o entendimento de GREGORIO (2022), para quem, os atores privados no ciberespaço exercem sua influência e isto está a desencadear preocupações sobre como estão estas entidades privadas desempenhando seu papel público:

> *[…] Public Powers still play a critical role in governing digital spaces and interfering with rights and freedoms. Nonetheless, the influence of private actors in the digital environment is increasingly raising concerns in terms of how these entities perform functions of public interest or, in some cases, mirror the exercise of public powers* (De Gregorio, 2022, p. 30).

Prossegue GIOVANNI DE GREGORIO advertindo que deixar questões desregulamentadas, como sistemas de IA, levaria a uma via aberta de tecnocracia ou *rule of tech,* nomeadamente pelas big tech, que teriam o poder de colocar os limites da proteção dos usuários de forma autônoma, numa escala global. Isso poderia refletir numa proteção deficitária ou aquém dos valores constitucionais consagrados em sistemas como da União Europeia e do Brasil.

> *Leaving algorithmic technologies without any democratic safeguard would lead to open the way to a form of techno-determinism, allowing not only public authorities but also private actors to govern algorithmic technologies to autonomously determine the standard of protection of rights and freedoms on a global scale* (De Gregorio, 2022, p. 286).

Logo, atores privados transnacionais consolidaram áreas de poderes delegadas e autônomas enquanto determinam, privadamente, os limites de

proteção de áreas sensíveis como moderação de conteúdo e proteção de dados. A ascensão do constitucionalismo digital, especialmente na Europa, também pode ser lida como uma reação contra o poder do plataformas on-line para definir seus valores em escala global de forma discricionária base.

Por isso, GIOVANNI DE GREGORIO (2022, p. 286) enxerga a autorregulação, em função do papel público dos atores privados na tecnologia, como fundamental para um quadro de respeito aos valores e princípios constitucionais, devendo estas instituições privadas, sobretudo, servirem como verdadeiros guardiões dos princípios democráticos, figurando como uma direção indicada pelo constitucionalismo digital:

> *Within this framework, the Union is going towards a different path. Rather than adopting a mere neoliberal approach or supporting the development of its model of the Internet, it is emerging at the intersection between the two models. The governance of values in the algorithmic society is not left either to private determinations through self-regulation or market intervention. The Union is consolidating a co-regulatory approach characterized by the definition of the value framework within which the private sector operates. Therefore, European constitutional values are not simply shaped by private determinations or by unaccountable forces, but are protected by a common regulatory framework injecting constitutional values in self-regulation. This result is not by chance but derives from the path of European digital constitutionalism* (De Gregorio, 2022, p. 290).

Desta maneira, o estabelecimento de respostas envolvendo a plêiade de atores públicos e privados, naquilo que se convencionou chamar de *governance* multinível, a partir de parâmetros constitucionais de obediência a regras e princípios comuns as democracias modernas, pautadas na dignidade da pessoa humana, somado à governação privada, parece-nos ser o caminho para uma regulação do espaço digital, especialmente a fim de transpassar os desafios limitadores de uma velha ordem de coisas, como ao excesso de poder das *big techs*, a extraterritorialidade destas e a soberania afeta às fronteiras dos países, desgastadas pela globalização e a velocidade de enfrentamento de novos problemas.

Ao constitucionalismo digital, portanto, cabe o papel de proteger o cidadão do abuso de poder, sendo imprescindível uma abordagem constitucional da relação entre utilizadores e empresas, o que significa regular o ciberespaço para que os direitos fundamentais sejam respeitados e promovidos e, neste ponto, uma abordagem que encare os problemas supracitados, parece-nos indicar uma maior participação privada no quadro regulatório, através de autorregulação e autorregulação regulada.

Pelo que, a União Europeia vem funcionando como um "farol" ao aprofundar o debate sobre a regulação e participação privada nesta, a fim de promover o respeito aos direitos fundamentais, dando indicações ao restante do mundo sobre suas experiencias regulatórias. A partir desta visão exposta, vamos adentrar a um estudo comparado entre Brasil e União Europeia, a fim de perceber como a regulação vem sendo realizada, especificamente em relação à proteção de dados e IA.

5.2) Quadro-geral regulatório comparativo da proteção de dados pessoais e inteligência Artificial entre Brasil (LGDP e PLIA) e União Europeia (RGPD e AIA)

Metodologicamente, optamos por recortar o quadro-geral regulatório a partir da legislação mais significativa para o tema dessa investigação que envolve proteção de dados pessoais e inteligência artificial, pelo que utilizamos a Lei Geral de Proteção de Dados brasileira (Lei 13.709/2018) e o PL n.º2338/23 (PLIA), que se encontra em estágio avançado de tramitação no Poder Legislativo brasileiro. Do ponto de vista da União Europeia, fizemos a análise partindo da Regulamento (UE) 2016/679, de 27 de abril de 2016 (RGPD), bem como do Regulamento para Inteligência Artificial, recentemente aprovado.

O espectro brasileiro regulatório do ciberespaço, apesar de andar mais lentamente, tem nítida inspiração europeia, como se nota da elaboração da LGPD e do PLIA, Entretanto, para LORDELO (2022), apesar ainda de não estar aprovado o PLIA, há um verdadeiro microssistema nacional de tutela dos direitos cibernéticos, composto pelo Código de Defesa do Consumidor (Lei 8.078/1990); Marco Civil da Internet (Lei 12.965/2014); Lei Geral de Proteção de Dados (Lei

13.709/2018); Lei do Processo Administrativo Federal (Lei 9.784/1999); Lei de Acesso à Informação (Lei 12.527/2011) e Resolução CNJ 332/2020.

Todavia, na última década, especialmente com a entrada em vigor da Lei Federal n.º12.965/2014, chamada de Marco Civil da Internet (MCI) que disciplina o uso da Internet no Brasil por meio da previsão de princípios, garantias, direitos e deveres para quem faz uso da rede, bem como da determinação de diretrizes para a atuação do Estado, houve uma proliferação de iniciativas de regulação do ciberespaço.

Como exemplo, em 2018 foi aprovada Lei Geral de Proteção de Dados (LGPD) - Lei Federal n.º13.709/18 – que dispôs sobre o tratamento de dados pessoais, inclusive nos meios digitais, por pessoa natural ou por pessoa jurídica de direito público ou privado, com o objetivo de proteger os direitos fundamentais de liberdade e de privacidade e o livre desenvolvimento da personalidade da pessoa natural, entrando em vigor somente 2 anos depois, notadamente em razão da pandemia de COVID-19, a partir de uma quase cópia do Regulamento Geral de Proteção de Dados da UE.

De forma estruturada, a LGPD estabeleceu um rol de direitos e princípios e fundamentos afetos à proteção de dados pessoais, que igualmente reflete os limites constitucionais consagrados na Constituição brasileira, como os direitos humanos, a dignidade da pessoa humana, a autodeterminação informativa, não discriminação, dentro outros.

Dentre os direitos previstos no artigo n.º18, destacamos o de acesso aos dados, correção de dados, anonimização, bloqueio ou eliminação de dados e de revogação do consentimento, pois guardam estreita ligação com a tutela de dados em sistemas de IA.

Em relação ao *governance* privado, em clara inspiração de autorregulação regulada, o artigo n.º50 designa que os controladores e operadores de dados poderão formular regras de boas práticas e de governança que estabeleçam as condições de organização, o regime de funcionamento, os procedimentos, incluindo reclamações e petições de titulares, as normas de segurança, os padrões técnicos, as obrigações específicas para os diversos envolvidos no tratamento, as ações educativas, os mecanismos internos de

supervisão e de mitigação de riscos e outros aspectos relacionados ao tratamento de dados pessoais.

Sendo assim, deve-se implementar programa de governança em privacidade que, no mínimo: i) demonstre o comprometimento do controlador em adotar processos e políticas internas que assegurem o cumprimento, de forma abrangente, de normas e boas práticas relativas à proteção de dados pessoais; ii) seja aplicável a todo o conjunto de dados pessoais que estejam sob seu controle, independentemente do modo como se realizou sua coleta; iii) seja adaptado à estrutura, à escala e ao volume de suas operações, bem como à sensibilidade dos dados tratados; iv) estabeleça políticas e salvaguardas adequadas com base em processo de avaliação sistemática de impactos e riscos à privacidade; v) tenha o objetivo de estabelecer relação de confiança com o titular, por meio de atuação transparente e que assegure mecanismos de participação do titular.

Por fim, para fiscalizar todo o cumprimento deste conjunto legal, determinou a criação da uma autoridade nacional de proteção de dados e de um conselho nacional de proteção de dados e privacidade.

Nesse ínterim, fora promulgada – com relativo atraso- a Emenda à Constituição n.º115, de 10 de fevereiro de 2022, que erigiu formalmente a proteção de dados como direito fundamental autônomo na Constituição Federal brasileira, assim como o fizera a Carta de Direitos Fundamentais da UE, no artigo n.8º, já em vigor há décadas.

O que atrai, sem questionamentos, o âmbito de proteção dos direitos e garantias fundamentais à lei de proteção de dados, deixando mais nítido as balizas que o legislador infraconstitucional, que a administração pública e, fundamentalmente, a iniciativa privada deverão observar, a fim de tutelar o direito em causa, em nítido diálogo multinível de inspiração oriunda do constitucionalismo digital.

Em relação à regulamentação da IA, o projeto de Lei Federal n.º2.323/2023 (PLIA), com tramitação iniciada a partir do Senado Federal tem um duplo objetivo, quais sejam, de um lado estabelecer direitos para proteção do elo mais vulnerável em questão, a pessoa natural que já é diariamente

impactada por sistemas de inteligência artificial, desde a recomendação de conteúdo e direcionamento de publicidade na Internet até a sua análise de elegibilidade para tomada de crédito e para determinadas políticas públicas e, de outro lado, dispor de ferramentas de governança e de um arranjo institucional de fiscalização e supervisão, cria condições de previsibilidade acerca da sua interpretação e, em última análise, segurança jurídica para inovação e o desenvolvimento tecnológico.

Conforme se lê da justificativa do PLIA (2023), a proposição parte da premissa, portanto, de que não há um *tradeoff* entre a proteção de direitos e liberdades fundamentais, da valorização do trabalho e da dignidade da pessoa humana face à ordem econômica e à criação de novas cadeias de valor. Pelo contrário, seus fundamentos e a sua base principiológica buscam tal harmonização, nos termos da Constituição Federal e o constitucionalismo digital parece-nos a ferramenta metodológica apta a legitimar esse diálogo.

Estruturalmente, a proposição estabelece uma regulação baseada em riscos (*risk-based approach*) e uma modelagem regulatória fundada em direitos(*right-based approach*). Apresenta ainda instrumentos de governança privada para uma adequada prestação de contas dos agentes econômicos desenvolvedores e utilizadores da inteligência artificial, incentivando uma atuação de boa-fé e um eficaz gerenciamento de riscos. O texto proposto, inicialmente, define fundamentos e princípios gerais para o desenvolvimento e utilização dos sistemas de inteligência artificial, que balizam todas as demais disposições específicas.

Em linha com os pilares de uma governação multinível proposta pelo constitucionalismo digital, o PLIA estabelece “fronteiras” seguras, em seu papel regulatório delimitador, já em seu artigo 2.º, ao determinar os fundamentos para desenvolvimento de sistemas de IA, tais como a centralidade da pessoa humana (antropocentrismo digital), respeito aos direitos humanos, valores democráticos, privacidade, proteção de dados, autodeterminação informacional, igualdade e não discriminação, dentre outros.

Dentre os direitos dos usuários de IA, previstos no artigo 5.º, ressalta-se o da proteção de dados pessoais, a privacidade, a não discriminação e de

supervisão humana em decisões tomadas por IA. Especificamente quanto as discriminações, as pessoas afetadas por decisões, previsões ou recomendações de sistemas de inteligência artificial têm direito a tratamento justo e isonômico, sendo vedadas a implementação e o uso de sistemas de inteligência artificial que possam acarretar discriminação direta, indireta, ilegal ou abusiva.

Ademais, o Capítulo IV traz previsões voltadas a governação dos sistemas de IA, reforçando que os agentes de inteligência artificial estabelecerão estruturas de governança e processos internos aptos a garantir a segurança dos sistemas e o atendimento dos direitos de pessoas afetadas, que incluirão, dentre outros: i) transparência quanto às medidas de governança adotadas no desenvolvimento e emprego do sistema de inteligência artificial pela organização; ii) medidas de gestão de dados adequadas para a mitigação e prevenção de potenciais vieses discriminatórios; iii) legitimação do tratamento de dados conforme a legislação de proteção de dados, inclusive por meio da adoção de medidas de privacidade desde a concepção e por padrão e da adoção de técnicas que minimizem o uso de dados pessoais.

Note-se neste ponto relativo à legitimação do tratamento de dados conforme a legislação de proteção de dados, há um claro diálogo de fontes legislativas, onde a proteção de dados pessoais e a legislação de IA deverão manter uma constante interação, mormente em função da simbiose entre dados e sistemas de IA, objetivando a construção de um ambiente digital seguro e juridicamente capaz de respeitar os direitos fundamentais dos usuários.

Todavia, se o sistema de IA for considerado de alto risco, medidas mais restritivas são impostas *ex lege*, como a realização de testes para avaliação de níveis apropriados de confiabilidade, conforme o setor e o tipo de aplicação do sistema de inteligência artificial, incluindo testes de robustez, acurácia, precisão e cobertura e medidas de gestão de dados para mitigar e prevenir vieses discriminatórios. Dessarte, a avaliação de impacto algorítmico de sistemas de inteligência artificial é obrigação dos agentes de inteligência artificial, sempre que o sistema for considerado como de alto risco pela avaliação preliminar, à teor do artigo n.º22.

Outrossim, o Capítulo VI traz as determinações para os códigos de boas práticas e de governança, de sorte que os agentes de inteligência artificial poderão, individualmente ou por meio de associações, formular códigos de boas práticas e de governança que estabeleçam as condições de organização, o regime de funcionamento, os procedimentos, inclusive sobre reclamações das pessoas afetadas, as normas de segurança, os padrões técnicos, as obrigações específicas para cada contexto de implementação, as ações educativas, os mecanismos internos de supervisão e de mitigação de riscos e as medidas de segurança técnicas e organizacionais apropriadas para a gestão dos riscos decorrentes da aplicação dos sistemas.

Especialmente em relação aos códigos de boas práticas de governação, reside o âmago da autorregulação, pois possibilita, para além da mera conformidade ao quadro-legal regulador imposto por determinado Estado, às empresas, por natureza multinacionais, uniformizar sua atuação, criando *standards* até mesmo mais protetivos dos direitos dos usuários, aumentando seu reconhecimento e valor de mercado, iniciando uma ciclo virtuoso no ciberespaço e forçando a concorrência a se adaptar, numa inspiração claramente de *governance* multinível.

Por fim, determina a PLIA a criação de uma Autoridade Competente para realizar a supervisão e a fiscalização que, para além de expedir normas regulamentares, deverá zelar pela proteção a direitos fundamentais e a demais direitos afetados pela utilização de sistemas de inteligência artificial; promover a elaboração, atualização e implementação da Estratégia Brasileira de Inteligência Artificial junto aos órgãos de competência correlata; promover e elaborar estudos sobre boas práticas no desenvolvimento e utilização de sistemas de inteligência artificial; e estimular a adoção de boas práticas, inclusive códigos de conduta, no desenvolvimento e utilização de sistemas de inteligência artificial.

Já no âmbito da União Europeia, esta vem assumindo a liderança e se revelando como um *farol* para o resto do mundo em termos de regulação do ambiente digital. Como é sabido, nos primórdios da internet nos Estados Unidos da América (EUA), por volta das décadas de 1970 e 1980, o desenho arquitetural era fortemente lastreado numa ideia liberal, quase anárquica, onde não cabia falar em qualquer tipo de regulamentação, sendo que a União Europeia primeiro

engendrou esforços na implantação da internet, para somente depois pensar em regulação. Assim, o professor espanhol MOISÉS BARRIO ANDRÉS explica-nos que:

> *La arquitectura técnica abierta, establecida a principios de los años sesenta y setenta, y la inexistencia de regulación de los años ochenta condujeron al auge del medio en los años noventa y a su ubicuidad en el siglo XXI. Esta circunstancia explica la tardanza de la Unión Europea de preocuparse primeiro por la implantación de internet en Europa, y luego respecto a su regulacion* (Andrés, 2020, p.95).

Portanto, a regulação jurídica da internet no espaço da União, tem um marco legislativo importante no ano de 1999, com a iniciativa "e-Europa: uma sociedade da informação para todos"[100], voltada a regular a inserção do espaço europeu no novo contexto da economia digital da sociedade da informação, através de objetivos claros, como a melhoria e ampliação de acesso à internet e o desenvolvimento do comércio eletrônico. Entretanto, vale ressalvar que desde 1996 a proteção de dados pessoais já possuía regramento via diretiva n.º96/9/CE, ainda que não vinculativo para Estados-Membros

Em sequência, destacamos as Diretivas 2000/31/CE do Parlamento Europeu e do Conselho, relativa a determinados aspectos jurídicos dos serviços na sociedade da informação, especialmente para o comércio eletrônico, bem como a Diretiva 1999/93/CE também do Parlamento Europeu e do Conselho, onde fora estabelecido um marco comunitário para a assinatura eletrônica, posteriormente derrogada pelo Regulamento n. º 910/2014.

Igualmente digno de atenção são os documentos *Agenda Digital Para Europa*, de 2010 e a *Estratégia para o Mercado Único Digital da Europa*, de 2015, onde há uma revisão da política digital para Europa, atualizando as exigências e desafios que as novas tecnologias impõem à sociedade e ao direito (Andrés, 2020).

[100] A passagem para uma economia digital baseada no conhecimento pode vir a ser um poderoso factor de crescimento, de competitividade e de criação de empregos, além de permitir melhorar a qualidade de vida dos cidadãos e o ambiente. Para criar esta " sociedade da informação para todos", a Comissão lançou, em 1999, a iniciativa eEurope, um programa ambicioso destinado a generalizar, tanto quanto possível, as tecnologias da informação. Disponível em https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=LEGISSUM:l24221. Acesso em 11 de julho de 2023.

Noutro giro, importante destacar a recente **"**Carta Portuguesa dos Direitos Fundamentais da Era Digital", que consignou um acervo de direitos, liberdades e garantias ajustáveis à *Sociedade Digital*:

> i) *Direito de Acesso ao Ambiente Digital; ii) Liberdade de Expressão e Criação; iii) Garantia do Acesso e Uso das Redes; iv) Direito à Protecção contra a Desinformação; v) Direito à Privacidade em Ambiente Digital; vi) Direito à Neutralidade da Internet; vii) Direito ao Desenvolvimento de Competências Digitais; viii) Direito à Identidade e Outros Direitos Pessoais; Direito ao Esquecimento; ix) Direitos em Plataformas Digitais; x) Direito à Ciber-segurança; xi) Direito à Liberdade de Criação e à Protecção dos Conteúdos; xii) Direito à Protecção contra a Geolocalização Abusiva; xiii) Direito ao Testamento Digital; xiv) Os Direitos dos Menores e sua específica protecção; xv) Direitos Digitais face à Administração Pública; xvi) Direito de Acção : o recurso à acção popular digital e a outras garantias*.

Deveras essencial também é a "Declaração sobre direitos e princípios digitais para a década digital[101]", promulgada pela União Europeia em 2022, que elenca seis pilares principiológicos nos quais devem se assentar o edifício regulatório digital a ser elaborado:

> 1) *Dar prioridade às pessoas no processo de transformação digital*
> 2) *Solidariedade e inclusão via: Conectividade; Educação, formação e competências digitais; Condições de trabalho justas e equitativas; e Serviços públicos digitais em linha*
> 3) *Liberdade de escolha em Interações com algoritmos e sistemas de inteligência artificial, num ambiente digital justo*
> 4) *Participação no espaço público digital;*
> 5) *Segurança, proteção e capacitação, via um ambiente digital protegido e seguro, com Privacidade e controlo individual dos dados e Proteção e capacitação das crianças e dos jovens no ambiente digital*
> 6) *Sustentabilidade*

A proteção de dados pessoais, como cerne da economia digital, naturalmente ganha relevância ímpar no Direito Público Digital. Por isso, a proteção de dados pessoais foi reconhecida como um direito fundamental pela

[101] Disponível em https://www.consilium.europa.eu/pt/press/press-releases/2022/12/15/declaration-on-digital-rights-and-principles-eu-values-and-citizens-at-the-centre-of-digital-transformation/. Acesso em 11 de julho de 2023.

Carta de Direitos Fundamentais da União Europeia, através do artigo 8º, n.1[102], bem como pelo artigo 16º, n 1[103], do Tratado sobre o Funcionamento da União Europeia (TFUE).

Igualmente, antes da entrada em vigor do Regulamento (EU) n.º 2016/679, de 27 de abril, comumente denominado de Regulamento Geral de Proteção de Dados (RGPD), a proteção de dados pessoais já era reconhecida através da Directiva 95/46/CE, posteriormente revogada em prol da utilização de um instrumento normativo dotado de vinculação obrigatória direta aos Estados-Membros- como é o caso do instituto jurídico do Regulamento- e capaz de uniformizar conceitos e procedimentos em matéria de dados pessoais em todo o espaço europeu[104].

Basicamente, o Regulamento (EU) n.º 2016/679é subdividido em onze capítulos, numa estrutura que se inicia pelo objeto e objetivo a ser tutelado, enumerando princípios e direitos do titulares em sequência, enquadrando o tratamento de dados e seus responsáveis, suas obrigações, direitos e responsabilidade, a questão da transferência internacional de dados, com uma posterior criação de estruturas de órgãos de fiscalização e formas de fazer cumprir as determinações, finalizando com os meios de direito de ação e reclamação dos lesados, com a consequente responsabilização e aplicação de sanções.

Assim, a teor do artigo 5.º da RGPD os princípios relativos ao tratamento de dados pessoais, são a licitude, a lealdade, e a transparência. Também são princípios a limitação da finalidade, a minimização dos dados e a exatidão, ao lado da limitação da conservação, integridade, confidencialidade e a responsabilidade. Já o postulado da transparência, previsto no art. 1. a) da

---

[102] *Todas as pessoas têm direito à protecção dos dados de carácter pessoal que lhes digam respeito.*

[103] *Todas as pessoas têm direito à proteção dos dados de caráter pessoal que lhes digam respeito.*

[104] Precisamente neste sentido é o considerando n.º10 do Regulamento que assim dispõe: *A fim de assegurar um nível de proteção coerente e elevado das pessoas singulares e eliminar os obstáculos à circulação de dados pessoais na União, o nível de proteção dos direitos e liberdades das pessoas singulares relativamente ao tratamento desses dados deverá ser equivalente em todos os Estados-Membros. É conveniente assegurar em toda a União a aplicação coerente e homogénea das regras de defesa dos direitos e das liberdades fundamentais das pessoas singulares no que diz respeito ao tratamento de dados pessoais. [...]*

RGPD, *in fine*, corrobora o tratamento lícito e equitativo dos dados pessoais, exigindo que a recolha, utilização consulta e tratamento deve ser transparente, sendo estas informações de fácil acesso e compreensão, exposta em linguagem clara e simples.

Ainda como imposição do referido princípio da transparência, é direito dos titulares serem alertados para os riscos, regras, garantias e direitos associados ao tratamento dos dados pessoais e para os meios de que dispõe para exercer seus direitos relativamente a esse tratamento.

Desta maneira, constata-se a regulação da União delimitando a atuação privada através da depuração de novos direitos a serem observados nas relações envolvendo o tratamento de dados pessoais, como ocorre entre usuários e plataformas digitais, como a transparência, a finalidade, a anonimização e a finalidade.

Outrossim, insta ressaltar a possibilidade de retirada do consentimento ou também chamado “direito de oposição”, que confere ao titular o direito de se opor ao tratamento de seus dados pessoais, como por exemplo para efeitos de comercialização direta, nos termos do art. 21.º, n.2. Com o avanço significativo da inteligência artificial, especialmente a generativa, o direito a oposição a decisões automatizadas mostra-se de especial relevância, haja vista que a prática operacional das plataformas necessita de amplo uso de dados pessoais e o uso de IA para executar os objetivos pelo que, em regra, o art. º22 da RGPD diz expressamente que:

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

Fazendo a devida ligação com o Regulamento sobre Inteligência Artificial, aprovado por maioria em 13 de março de 2024, a UE chegou a um acordo sobre o tema, consolidando o texto sobre a regulação de sistemas de IA. Representando o regulamento sobre IA como pioneiro e indicativo para o restante dos ordenamentos jurídicos ao redor do globo, estabelecendo um caminho claro para um desenvolvimento seguro e centrado no ser humano em

termos da IA, ou seja, tutelando direitos a partir da dignidade humana e proteção de direitos fundamentais.

Por isso, o objetivo da regulamentação são de quatro vieses, quais sejam: i) garantir que os sistemas de IA colocados no mercado da União e utilizados sejam seguros e respeitem a legislação em vigor em matéria de direitos fundamentais e valores da União; ii) garantir a segurança jurídica para facilitar os investimentos e a inovação no domínio da IA; iii) melhorar a governação e a aplicação efetiva da legislação em vigor em matéria de direitos fundamentais e dos requisitos de segurança aplicáveis aos sistemas de IA; iv) facilitar o desenvolvimento de um mercado único para as aplicações de IA legítimas, seguras e de confiança e evitar a fragmentação do mercado.

Consoante se lê das razões e motivos da proposta, o legislador europeu decidiu construir essa proposta a partir destes quatro pilares, característicos de uma abordagem horizontal, equilibrada e proporcional, onde há uma limitação aos requisitos mínimos necessários para dar resposta aos riscos e aos problemas associados à IA, sem restringir ou prejudicar indevidamente a evolução tecnológica ou aumentar desproporcionalmente o custo de colocação no mercado das soluções de IA:

> A proposta estabelece um quadro jurídico sólido e flexível. Por um lado, as suas escolhas regulamentares fundamentais, incluindo os requisitos baseados em princípios que os sistemas de IA devem respeitar, são abrangentes e estão preparadas para o futuro. Por outro lado, cria um sistema regulamentar proporcionado, centrado numa abordagem regulamentar baseada no risco bem definida que não cria restrições desnecessárias ao comércio e na qual a intervenção jurídica é adaptada às situações concretas em que existe um motivo de preocupação justificado ou em que tal preocupação pode ser razoavelmente antecipada num futuro próximo. Ao mesmo tempo, o quadro jurídico inclui mecanismos flexíveis que permitem a sua adaptação dinâmica à medida que a tecnologia evolui e surgem novas situações preocupantes.

Nota-se que o *mens legis* está umbilicalmente atrelado aos problemas enfrentados pela clássica regulamentação – já exposto acima, como velocidade, territorialidade e soberania-, trazendo os atores privados para o centro da regulação do ciberespaço, visando suprir estas lacunas, prestigiando a

governação multinível, em especial a autorregulamentação setorial das plataformas digitais.

De mais, servindo de inspiração para a legislação brasileira, o AIA estabelece regras harmonizadas para o desenvolvimento, a colocação no mercado e a utilização de sistemas de IA na União na sequência de uma abordagem proporcionada baseada no risco (*risk-based approach*), com condutas já proibidas *a priori*, consoante se lê do artigo n.5º, como por exemplo:

> 1.a) A colocação no mercado, a colocação em serviço ou a utilização de um sistema de IA que empregue técnicas subliminares que contornem a consciência de uma pessoa para distorcer substancialmente o seu comportamento de uma forma que cause ou seja suscetível de causar danos físicos ou psicológicos a essa ou a outra pessoa (Regulamento para Inteligência Artificial (AIA)).

Ou seja, existem requisitos específicos para sistemas de IA de risco elevado e obrigações para os operadores desses sistemas, considerando como risco elevado quando o sistema de IA destina-se a ser utilizado como um componente de segurança de um produto ou é, ele próprio, um produto abrangido pela legislação de harmonização da União.

Como previsto também no PLIA brasileiro, para esses sistemas de IA de risco elevando deve-se cumprir obrigações mais severas, descritas no “Capitulo 2” da proposta, tais como criação e implementação de gestão de riscos, governação de dados, documentação técnica, manutenção de registos, transparência, prestação de informações aos utilizadores, cibersegurança e a supervisão humana, em nítida inspiração de um constitucionalismo digital.

À evidência, o texto aprovado sobre regulamentação de IA é claro ao dispor que está em consonância com Regulamento de Governação de Dados, à Diretiva Dados Abertos e a outras iniciativas estabelecidas na Estratégia europeia para os dados, que criarão mecanismos e serviços de confiança para a reutilização, a partilha e o agrupamento de dados, elementos essenciais para o desenvolvimento de modelos de IA baseados em dados de elevada qualidade.

Vale destacar que o Regulamento de Governação de Dados é a primeira de um conjunto de medidas levado à cabo pela Estratégia Europeia para

os Dados de 2020, ambicionando promover a disponibilização de dados para serem utilizados, aumentando a confiança nos intermediários de dados e reforçando os mecanismos de partilha de dados em toda a UE. O instrumento aborda as seguintes questões como i) disponibilização de dados do setor público para reutilização, em situações em que esses dados estejam sujeitos a direitos de terceiros; ii) a partilha de dados entre empresas, mediante remuneração, independentemente da forma que assuma; iii) a autorização da utilização de dados pessoais através de um intermediário de partilha de dados pessoais, concebido para ajudar as pessoas singulares a exercerem os seus direitos ao abrigo do Regulamento Geral sobre a Proteção de Dados (RGPD); e iv) a autorização da utilização de dados com finalidades altruístas.

Em termos de governação, o Regulamento sobre Inteligência Artificial da União Europeia, o capítulo 2 estabelece os requisitos legais aplicáveis aos sistemas de IA de risco elevado relativamente aos dados e à governação de dados, à documentação e à manutenção de registos, à transparência e à prestação de informações aos utilizadores, à supervisão humana, à solidez, à exatidão e à segurança, enquanto o capítulo 3 indica às obrigações acrescidas aos que operam sistemas de IA de alto risco, como a criação, implementação e documentação de gestão de riscos.[105]

Ademais, os sistemas de IA de risco elevado que utilizem técnicas que envolvam o treino de modelos com dados devem ser desenvolvidos com base em conjuntos de dados de treino, validação e teste que cumpram requisitos, como prévia avaliação, detecção de vieses e deficiência de dados, buscando qualificar o conjunto de dados utilizados, para minimizar efeitos indesejáveis e ilegais dos resultados dos sistemas de IA, fugindo à lógica do *garbage in, garbage out.*

[105] Deverá compreender: i) identificação e análise dos riscos conhecidos e previsíveis associados a cada sistema de IA de risco elevado; ii) estimativa e avaliação de riscos que podem surgir quando o sistema de IA de risco elevado é usado em conformidade com a sua finalidade prevista e em condições de utilização indevida razoavelmente previsíveis; iii) avaliação de outros riscos que possam surgir, baseada na análise dos dados recolhidos a partir do sistema de acompanhamento pós-comercialização e; iv) adoção de medidas de gestão de riscos adequadas em conformidade com o disposto nos números que se seguem (cf. Artigo 9º, n.2)

No sentido de fixar os parâmetros sobre a regulação dos sistemas de IA de alto risco na UE, o AIA ainda determina a i) manutenção de registros; ii) transparência e prestação de informações aos utilizadores; iii) a supervisão humana; iv) exatidão, solidez e cibersegurança.

Quanto aos fornecedores de sistemas de IA de risco elevado, estes devem observar, dentre outros, i) implantação de gestão da qualidade; ii) elaboração de documentação técnica; iii) avaliação de conformidade; iv) registros gerados automaticamente; v) medidas corretivas; vi) dever de informação e cooperação.

Em relação aos códigos de conduta, o Título IX estabelece um quadro para a criação de códigos de conduta, que visa incentivar os fornecedores de sistemas de IA que não são de risco elevado a aplicar voluntariamente os requisitos obrigatórios aplicáveis aos sistemas de IA de risco elevado, elevando a proteção dos utilizadores – e do próprio ciberespaço- de maneira voluntária.

Os fornecedores de sistemas de IA que não são de risco elevado podem criar e aplicar autonomamente os códigos de conduta. Esses códigos também podem incluir compromissos voluntários relacionados, por exemplo, com a sustentabilidade ambiental, a acessibilidade das pessoas com deficiência, a participação das partes interessadas na conceção e no desenvolvimento de sistemas de IA e a diversidade das equipas de desenvolvimento, erigindo o postulado da devida diligência empresarial no universo tecnológico.

Por tudo isso, parece-nos existir uma regulação via AIA inspirada no *governance* multinível, chamando o particular a participar ativamente do quadro-legal regulatório, contemplando a maximização e precisão das respostas aos problemas existentes e ainda por vir, a partir de uma leitura antropocêntrica digital.

De mais a mais, resta finalizar a abordagem europeia de regulação do ciberespaço ressaltando os diplomas do *Digital Service Act* (DSA) e *Digital Market Act* (DMA), que previu regras regulatórias designadamente para os marketplaces e plataformas digitais em linha, com o objetivo de assegurar um ambiente em linha seguro, previsível e fiável, combatendo a difusão de conteúdos ilegais em linha e os riscos sociais que a difusão de desinformação

ou de outros conteúdos pode gerar, e no qual os direitos fundamentais consagrados na Carta sejam eficazmente protegidos e a inovação seja facilitada.

Para TORNADA (2023) o recente Regulamento de Serviços Digitais (DSA) tem o potencial de ser revolucionário em termos de proteção dos direitos fundamentais no ciberespaço, pois tutela os perigos de priorizar algoritmicamente o conteúdo de acordo com os interesses dos usuários e comportamento, nomeadamente como catalisador dos riscos sistémicos de disseminação de desinformação e conteúdo ilegal.

Para tanto, observamos a aplicação do constitucionalismo digital multinível no referido diploma, na medida em que o artigo n.º14[106], 4, que regulamenta os termos e condições das plataformas digitais diz, expressamente, que devem agir levando em conta os direitos e interesses legítimos das partes, incluindo os direitos fundamentais dos destinatários dos serviços, como consagrado na Carta de Direitos Fundamentais da UE.

Isso quer dizer, as plataformas digitais em sua relação com os usuários, deve ter em conta, aplicar e zelar pelos direitos fundamentais da Carta da UE, como a observância do direito à igualdade, a liberdade de expressão, a autonomia e ao devido processo legal. Significa, na prática, que ao realizar, especialmente, moderação e recomendação de conteúdos, suspensão ou exclusão de perfis e imposição de penalidades deverá haver a observância imperiosa de direitos fundamentais dos usuários para além dos códigos de conduta e termos e condições de serviços, como o contraditório, ampla defesa, recurso, direito de produzir provas e decisão racionalmente justificável à luz do direito.

Isso significa, portanto, a demonstração clara de *normatividade* multinível, típico do constitucionalismo digital, para que os particulares observem

[106] Os prestadores de serviços intermediários agem de forma diligente, objetiva e proporcionada na aplicação e execução das restrições referidas no n. 1, tendo devidamente em conta os direitos e interesses legítimos de todas as partes envolvidas, incluindo os direitos fundamentais dos destinatários do serviço, como a liberdade de expressão, a liberdade e o pluralismo dos meios de comunicação social e outros direitos e liberdades fundamentais, tal como consagrados na Carta.

os direitos fundamentais reconhecidos no âmbito da União, especialmente na relação horizontal travada num contrato de prestação de serviços.

À evidência, o que se extrai deste panorama, é a vontade política da UE em regular os assuntos voltados à internet. Para isso, vem adotando uma política consistente num conjunto formal de legislações e *soft law* -como as declarações suprarreferidas. Quer dizer, esse *legal framework* acaba por indicar e corroborar a existência de uma esfera pública digital, seus limites e domínios e, por consequência, de um direito público digital correspondente a ser tutelado.

CONCLUSÃO:

É certo dizer que os avanços da tecnologia, em especial aquelas relacionadas à Internet, criaram um novo espaço para o debate de ideias e a formação da opinião pública que, em razão do alcance quase ilimitado de pessoas e lugares, passou a ser a nova *locus* pública dos tempos atuais: as plataformas digitais.

De um tempo onde não havia nenhuma, ou quase nenhuma, regulação, experimentamos abusos, vigilância, discriminações e acumulo de poder inigualável pelo setor privado, especialmente as grandes empresas de tecnologia, levando à sociedade a debater a regulação do ciberespaço.

Todavia, a disrupção foi tamanha que o direito, nomeadamente, o Constitucional, com suas ferramentas tradicionais, pautadas nas noções soberania, legiferação morosa, fronteiras delimitadas e bem definas, já não se ajustavam de forma hermética aos problemas e desafios atuais, pois a internet desconhece fronteiras e não se sujeita a um órgão central mundial capaz de impor regras e vivencia mudanças imediatas, em velocidade instantânea.

Nesse momento, surge uma tese denominada de Constitucionalismo Digital, que estuda a compreensão destes fenômenos e oferta respostas aos desafios impostos pela nova esfera pública digital, em especial a tutela de direitos fundamentais on-line e as formas de governação e regulação global dos problemas.

Somado a este estado de coisas, temos o exponencial crescimento da inteligência artificial, que vale de enorme quantidade de dados, em especial os de natureza pessoal, para seu desenvolvimento e utilização. A preocupação com o futuro da sociedade diante do avanço desta tecnologia, particularmente a chamada inteligência artificial generativa, que tem a capacidade de criar novas informações, leva a um acréscimo do debate sobre a regulação do espaço digital, em vista da possibilidade de a máquina, em breve, ganhar consciência e passar a pensar livremente como um ser humano.

Vislumbramos um esforço regulatório importante partindo da União Europeia, com o pioneirismo em tutelar em diversos temas caros à internet, como a própria proteção de dados pessoais, mercados e serviços digitais e a inteligência artificial, servindo de farol e inspiração para o resto do mundo.

Percebemos, entretanto, que o constitucionalismo digital oferta como resposta um *governance* multinível como solução aos problemas de regulação clássico, como adstrição ao território, soberania e velocidade de resposta. Ou seja, a regulação deve compreender a autorregulação também, chamando os atores para responsabilidade de promoverem sua regulamentação interna em padrões mundialmente aceitáveis, pois dotada na grande maior parte de velocidade e precisão, mas desde que respeitado o quadro-geral limitador imposto pelo Estado de Direito, como o respeito aos direitos fundamentais e aos princípios democráticos.

Assim, notamos que nos diplomas analisados sobre a proteção de dados pessoais e inteligência artificial, tanto na União Europeia, quanto no Brasil, há um nítido componente de governação multinível, com protagonismo do particular em torno do escopo regulatório, com direitos e deveres, numa relação de simbiose com o poder público regulador ou, ao menos, de uma verticalidade atenuada, o que corrobora a tese inicial de que a proteção de dados pessoais é vital para inteligência artificial e a regulação caminha em direção ao constitucionalismo digital.

REFERÊNCIAS BIBLIOGRÁFICAS

ANDRÉS, Moisés Barrio. **Manuel de Derecho Digital**. 1. ed. Valência: Tirant to Blanch, 2020.

ARBIX, Daniel. A Importância da Privacidade por Design e por Default (Privacy By Design and By Default). In: CUEVA, Ricardo Villas Bôas.; DONEDA, Danilo; MENDES, Laura. Schertel. **Lei Geral de Proteção de Dados (Lei nº13.709/2018:** a caminho da efetividade: contribuições para implementação da LGPD. São Paulo: Thomson Reuters Brasil, 2020.

BALKIN, Jack. Free Speech in the Algorithmic Society: Big Data, Private Governance and New School Speech Regulation. **Yale Law School Public Law Research Paper**, 09 set. 2017.

BIONI, Bruno. **Proteção de dados pessoais:** a função e os limites do consentimento. 3. ed. Rio de Janeiro: Forense, 2021.

BUSCH, Kristen. 2023. Generative Artificial Intelligence and Data Privacy: A Primer. **Congressional Research Service**. May 23, 2023.

CAMPOS, Ricardo. **Metamorfoses do Direito Global:** sobre a interação do direito, tecnologia e tempo. São Paulo: Contracorrente, 2022.

CASTELLS, Manuel. **A Sociedade em Rede**. Tradução de Roneide Venâncio Majer. 6ª. ed. São Paulo: Paz e Terra, 1999.

CASTRO, Raquel Brízida. **Direito Constitucional:** Ciberspeaço e Tecnologia Declínio do Constitucionalismo na UE? Coimbra: Almedina, 2023.

CELESTE, Edoardo. Digital Constitutionalism: a new systematic theorisation. **International Review of Law, Computers & Technology**, v.33, n.1, 2019.

DANAHER, John. **The Threat of Algocracy**: Reality, Resistance and Accommodation; Philosophy & Technology, 2016.

DE GREGORIO, Giovanni. **Digital Constitucionalism in Europe:** reframing rights and powers in the algorithmic society. London: Cambridge University Press, 2022.

—. The normative power of Artificial Intelligence : Working Paper n.º4/2023. **Católica Global of Law**, 2023.

DONEDA, Danilo. **Da Privacidade à proteção de dados pessoais:** elementos da formação da Lei Geral de proteção de dados. São Paulo: Thomson Reuters Brasil, 2019.

FARINHO, Domingos.; MÜLLER DORNELAS, Felipe. O controle da aplicação de inovações tecnológicas no Direito: o caso da utilização de algorítmos para recomendar e moderar conteúdos digitais em plataformas no recente Regulamento dos Serviços Digitais (Digital Service Act) da União Europeia. **no prelo**, 2024

GILL, Lex.; REDEKER, Dennis.; GASSER, Urs. Towards Digital Constitutionalism? Mapping Attempts to Craft an Internet Bill of Rights. **Berkman Center Research**, Cambridge, 09 nov. 2015.

HACKER, Philipp; CORDES, Johann; ROCHON, Janina. **Regulating Gatekeeper AI and Data:** Transparency, Access, and Fairness under the DMA, the GDPR, and beyond, 2023.

HACKER, Philipp; NOVELLI, Cláudio; CASOLARI, Federico; SPEDICATO, Giorgio; FLORIDI, Luciano. **Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity**, 14 de jan, 2024.
HACKER, Philipp. **AI Regulation in Europe**, 07 mai, 2020.

—. Sustainable AI Regulation,01 jun, 2023.

—. Expert Testimony on Generative AI. **Deutscher Bundestag**, 23 mai, 2023.

—. A Legal Framework for AI Training Data- From First Principles to the Artificial Intelligence Act. **Law, Innovation and Technology**. 2021

HOFFMANN-RIEM, Wolfgang. **Teoria Geral do Direito Digital:** transformação digital desafios para o direito. Tradução de Italo Fuhrmann. Rio de Janeiro: Forense, 2021.

KUNER, Christopher; CATE, Fred H.; LYNSKEY, Orla; MILLARD, Christopher; LOIDEAIN, Nora Ni; SVANTESSON, Dan Jerker B. Expanding the artificial intelligence-data protection debate. **International Data Privacy Law**, v. 8, n. 4, p. 289–292, nov. 2018.

LORDELO, João Paulo. **Constitucionalismo digital e devido processo legal**. São Paulo: JusPodivm, 2022.

MACHADO, Diego. **Algorítmos e proteção de dados pessoais:** tutela de direitos na era dos perfis. São Paulo: Almedina Brasil, 2023.

MATIUZZO, Marcela. Discriminação Algorítmica: reflexos no contexto da Lei Geral de Proteção de Dados Pessoais. In: CUEVA, Ricardo Villas Bôas.; DONEDA, Danilo.; MENDES, Laura Schertel. **Lei Geral de Proteção de Dados (Lei 13.709/2018):** a caminho da efetividade: contribuições para implementação da LGPD. São Paulo: Thomson Reuters Brasil, 2020.

MCCARTHY, John. What is Artificial Intelligence. **Stanford University**, Stanford, 12 nov. 2007.

MENDES, Gilmar Ferreira.; BRANCO, Paulo Gonet. **Curso de Direito Constitucional**. 12º. ed. São Paulo: Saraiva, 2017.

MENDES, Gilmar Ferreira.; FERNANDES, Victor Oliveira. Constitucionalismo digital e jurisdição constitucional: uma agenda de pesquisa para o caso brasileiro. **Revista Brasileira de Direito**, Passo Fundo, 16, n. 1, out. 2020.

MENDES, Laura Schertel. **Privacidade, proteção de dados e defesa do consumidor:** linhas gerais de um novo direito fundamental. São Paulo: Saraiva, 2014.

MENDES, Laura Schertel. *et al*. A discriminação algorítmica à luz da Lei Geral de Proteção de Dados. In: MENDES, L. S., *et al*. **Tratado de Proteção de Dados Pessoais**. Rio de Janeiro: Forense, 2021.

MENDES, Laura Schertel. Autodeterminação informacional: origem e desenvolvimento conceitual na jurisprudência da Corte Constitucional Alemã. *In*: CUEVA, Ricardo Villas Bôas; DONEDA, Danilo; MENDES, Laura Schertel (coord.) **Lei Geral de Proteção de Dados (Lei 13.709/2018):** a caminho da efetividade. Contribuições para implementação da LGPD. São Paulo: Thomson Reuters Brasil, 2020.

MIRANDA, Jorge. **Direitos Fundamentais 2ªEd**. Coimbra: Almedina, 2018.

MORAIS, Carlos Blanco de. **O sistema político no contexto da erosão da democracia representativa**. Coimbra: Almedina, 2020.

MÜLLER DORNELAS, Felipe. O direito ao esquecimento e dignidade da pessoa humana e a crítica necessária à tese fixada no caso aida curi - recurso extraordinário 1.010.606 do Supremo Tribunal Federal. **International Journal ow Law Jus Scriptum**, Lisboa, 2022.

NIKOLENKO, Sergei I. Synthetic Data for Deep Learning. **Computer Science, Computer Science (R0). Springer Optimization and Its Applications**, Springer Cham, 26, set, 2019.

NOVAIS, Jorge Reis. **A dignidade da pessoa humana:** Dignidade e Direitos Fundamentais. Coimbra: Almedina, 2018.

—. Direitos Fundamentais e Justiça Constitucional. Lisboa: AAFDL, 2019.

—. Princípios Estruturantes do Estado de Direito. Coimbra: Almedina, 2019.

PAESANI, Liliana Minardi. A publicidade móvel e a vulnerabilidade do consumidor. *In*: MORATO, Antônio Carlos; NERI, Paulo de Tarso (org**). 20 anos do Código de Defesa do Consumidor: estudos em homenagem ao professor José Geraldo Brito**. São Paulo: Atlas, 2010.

POLLICINO, Oreste. Judicial protection of fundamental rights in the transition from the world of atoms to the world of bits: the case of freedom of speech. **European Law Journal**, v.25, 2019.

POLLICINO, Oreste; DE GREGORIO, Giovanni. **Constitutional Law in the Algorithmic Society**. Cambridge: Cambridge University Press, 2021

—. **A Constitutional-Driven Change of Heart ISP Liability and Artificial Intelligence in the Digital Single Market.** The Global Community Yearbook of International Law and Jurisprudence, 2019.

RAMOS, Anatália Saraiva Martins. **Generative Artificial Intelligence based on large language models - tools for use in academic research**. In *SciELO Preprints*. 2023.

REINO UNIDO. ***Safety and Security Risks of Generative Artificial Intelligence to 2025****. Londres, 2023.* Disponível em: <https://assets.publishing.service.gov.uk/media/653932db80884d0013f71b15/generative-ai-safety-security-risks-2025-annex-b.pdf>.

RODOTÀ, Stefano. **A Vida na Sociedade da Vigilância:** a privacidade hoje. Tradução de Danilo Doneda e Luciana Cabral. Rio de Janeiro: Renovar, 2008.

SCHUILENBURG, Marc.; PEETERS, Rik. **The Algorithmic Society:** Technology, Power and Knowledge. 1. ed. New York: Routledge, 2020.

SCHWAB, Klaus. **A quarta revolução industrial**. São Paulo: EdiPro, 2016.

SOUSA, Simão Mendes de. **Constitucionalismo Digital:** uma introdução. Coimbra: Almedina, 2022.

TAULLI, Tom. **Artificial Intelligence Bascis:** a Non-Technical Introduction. Los Angeles: Apress Springer , 2019.

TEUBNER, Gunther. **Fragmentos Constitucionais:** constitucionalismo social na globalização. 2ª. ed. São Paulo: Saraiva, 2020.

TORNADA, João. How (not) to deal with the "bubble effect" in cyberspace: the case of the EU Digital Service Act. **Brooklyn Journal of International Law**, New York, 30 dez. 2023. 97-129.

VESTING, Thomas. The Impact of Artificial Intelligence on the Structures of the Modern Public Sphere. In **The Rule of Law in Cyberspace. Law, Governance and Technology Series**. Switzerland: Springer, v. 49ª, 2021.

WEBER, Max. **Sociologia**. 7. ed. São Paulo: Atica, v. 5, 2003.

XAVIER, Fábio Correa. **Alucinações de IA: o lado perverso da criatividade**.2023.

ZUBOFF, Shoshana. **A era do Capitalismo de Vigilância:** a disputa por um futuro humano na nova fronteira do poder. Tradução de Luís Filipe Silva e Miguel Serras Pereira. Lisboa: Relógio D'Água, 2020.

Legislação e Jurisprudência:

BRASIL. Autoridade Nacional de Proteção de Dados (ANDP). **Análise preliminar do Projeto de Lei nº 2338/2023, que dispõe sobre o uso da Inteligência Artificial.** Brasília, 2023. Disponível em: <https://www.gov.br/anpd/pt-br/assuntos/noticias/anpd-publica-analise-preliminar-do-projeto-de-lei-no-2338-2023-que-dispoe-sobre-o-uso-da-inteligencia-artificial>.

—. **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Presidência da República, 2023. Disponível em: <https://www.planalto.gov.br/ccivil_03/constituicao/constituicaocompilado.htm>.

—.**Decreto-Lei n.º4.657, de 4 de setembro de 1942**. Dispõe sobre a Introdução Às Normas do Direito Brasileiro. Brasília, 2023. Disponível em: <https://www.planalto.gov.br/ccivil_03/decreto-lei/del4657.htm>.

—. Senado Federal. **Projeto de Lei n.º2.338, de 2023 Dispõe sobre o uso da Inteligência Artificial.** Brasília: Senado Federal, 2023. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>.

—.Senado Federal. **Relatório final da comissão de juristas responsável por subsidiar elaboração de substitutivo sobre inteligência artificial no Brasil**. Brasília: Senado Federal, 2022. Disponível em: <https://www12.senado.leg.br/ecidadania/visualizacaoaudiencia?id=26627>.

—. Supremo Tribunal Federal. **ADI nº 6388/DF**. Relatora: Ministra Rosa Weber. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=754357772>.

—. Supremo Tribunal Federal.**RE n.º1.010.606/RJ**. Relator: Ministro Dias Toffoli. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=755910773>.

—. Lei n.º13.709, de 14 de agosto de 2018. **Lei Geral de Proteção de Dados Pessoais (LGPD).** Disponível em: <https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>.

UNIÃO EUROPEIA. **Carta dos Direitos Fundamentais da União Europeia** (2016/C 202/02), de 18 de dezembro de 2000. Disponível em : <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:12016P/TXT&from=FR>.

—. **CONSELHO DA EUROPA**. Draft framework convention on artificial intelligence, human rights, democracy and the rule of law. Comitê para Inteligência Artificial. 2023. Disponível em: <https://rm.coe.int/cai-2023-28-draft-framework-convention/1680ade043>.

—. **Declaração Europeia sobre os Direitos e Princípios Digitais** para a Década Digital. Disponível em: <https://digital-strategy.ec.europa.eu/pt/policies/digital-principles>.

—.**Diretiva 96/9/CE** do Parlamento Europeu e do Conselho relativa à protecção jurídica das bases de dados de 11 de Março de 1996. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX%3A31996L0009>.

—.**Diretiva 2000/31/CE** do Parlamento Europeu e do Conselho relativa a certos aspectos legais dos serviços da sociedade de informação, em especial do comércio electrónico, no mercado interno, de 8 de Junho de 2000. Disponível em : <https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=celex%3A32000L0031>.

—. **Grupo Independente de Peritos de alto nível sobre inteligência artificial**. Uma definição de IA: principais capacidades e disciplinas científicas, 2018. Disponível em: <https://digital-strategy.ec.europa.eu/pt/policies/expert-group-ai>.

—.**Regulamento (UE) 2022/2065** do Parlamento Europeu e do Conselho relativo a um mercado único para os serviços digitais e que altera a Diretiva 2000/31/CE (Regulamento dos Serviços Digitais), de 19 de outubro de 2022. Disponível em <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:32022R2065>.

—.**Regulamento (UE) 2022/2065** do Parlamento Europeu e do Conselho relativo à disputabilidade e equidade dos mercados no setor digital e que altera

as Diretivas (UE) 2019/1937 e (UE) 2020/1828 (Regulamento dos Mercados Digitais), de 14 de setembro de 2022. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=celex:32022R1925>.

—.**Regulamento (UE) 2016/679** do Parlamento Europeu e do Conselho relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados), de 27 de abril de 2016. Disponível em: <https://eur-lex.europa.eu/eli/reg/2016/679/oj>.

—. **Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial** (regulamento inteligência artificial) e altera determinados atos legislativos da união, de 21 de abril de 2021. Disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52021PC0206.

—.O Regulamento (UE) 2023/2854 do Parlamento Europeu e do Conselho relativo a regras harmonizadas sobre o acesso equitativo aos dados e a sua utilização, de 13 de dezembro de 2023. Disponível em : https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX:32023R2854

—.**Tribunal de Justiça da União Europeia**. Decisão n.º CJEU, C-252/21, *Meta vs. Bundeskartellamt*, ECLI:EU:C:2023:537. 2021. Disponível em: <https://curia.europa.eu/juris/documents.jsf?num=C-252/21>.

—.**Tribunal de Justiça da União Europeia**. CJEU, C-634/21, QG vs. SCHUFA, ECLI:EU:C:2023:957, para. 73. 2021. Disponível em: <https://curia.europa.eu/juris/documents.jsf?num=C-634/21>.

PORTUGAL**. Constituição da República Portuguesa**, de 2 de abril de 1976. Lisboa. Disponível em : <https://www.parlamento.pt/Legislacao/Paginas/ConstituicaoRepublicaPortuguesa.aspx>.

—. **Lei n.º 27/2021 Carta Portuguesa de Direitos Humanos na Era Digital**, de 17 de maio. Lisboa. Disponível em : <https://diariodarepublica.pt/dr/legislacao-consolidada/lei/2021-164870244>.