# A PROPOSTA REGULATÓRIA DA UNIÃO EUROPEIA PARA A *INTELIGÊNCIA ARTIFICIAL* (3ª. parte) – Sistemas de "alto risco"

**Demócrito Reinaldo Filho**
Desembargador do TJPE

A *Comissão Europeia*, braço executivo da União Europeia, apresentou no dia 21 de abril deste ano a sua proposta para regulamentação das tecnologias de *inteligência artificial* (IA)[1]. A proposta, que recebeu o nome de *Artificial Intelligence Act*[2], tem uma abordagem regulatória calcada na **hierarquização dos riscos** oferecidos por sistemas e tecnologias que usam IA[3].

Segundo essa visão regulatória baseada nos riscos (*risk-based regulatory approach*), as restrições e exigências aumentam conforme maiores sejam os riscos que os sistemas de IA possam oferecer a direitos e garantias

---

[1] **Inteligência artificial** (por vezes mencionada pela sigla em português **IA** ou pela sigla em inglês **AI** - *artificial intelligence*) é a inteligência similar à humana exibida por sistemas de *software*, além de também ser um campo de estudo acadêmico.

[2] O nome completo atribuído à proposta legislativa é: *Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS*. Cópia em PDF da proposta de regulação pode ser obtida em: https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence-artificial-intelligence

[3] Para melhor compreensão do modelo regulatório baseado nos riscos oferecidos pelos sistemas de IA, sugerimos a leitura de nosso primeiro artigo sobre a proposta de Regulamento da Comissão Europeia. Ele serve de introdução ao tema da regulação da inteligência artificial na UE e é a primeira parte de nossa análise sobre o Regulamento. Sua leitura é fundamental para a compreensão do escopo e objetivo do Regulamento. O titulo do primeiro artigo é: *A PROPOSTA REGULATÓRIA DA UNIÃO EUROPEIA PARA A INTELIGÊNCIA ARTIFICIAL (1ª. parte) – A hierarquização dos riscos*. Acessível em: https://jus.com.br/artigos/90816.

fundamentais dos indivíduos. Os níveis de regulação são diferentes de acordo com os riscos, variam conforme os riscos que os sistemas de IA possam apresentar a valores da sociedade e direitos das pessoas.

Nessa acepção, o regulamento classifica os sistemas de IA em três diferentes patamares de risco: os de "risco inaceitável" (*unacceptable risk*), os de "risco elevado" (*high-risk*) e os de "risco limitado" (*limited risk*) ou de "risco mínimo" (*minimal risk*). O desenvolvimento e utilização de sistemas que apresentem "risco inaceitável" são completamente vedados, em razão do elevado potencial de vulneração de direitos fundamentais.

Em artigo anterior[4], descrevemos os sistemas que se enquadram na categoria de "risco inaceitável" e, por isso, são banidos do mercado, por configurarem práticas de inteligência artificial intoleráveis, violadoras de direitos fundamentais das pessoais. No presente artigo, trataremos de examinar as características dos sistemas de "alto risco" e os requisitos para que possam ser livremente comercializados.

**Sistemas de "alto risco"**

No Título III, o Regulamento contém regras para sistemas de IA que criam alto risco para a saúde, segurança e direitos fundamentais das pessoas. Seguindo a concepção baseada no risco (*risk-based approach*), essa parte do regulamento descreve sistemas que não são proibidos, mas que sofrem severas restrições quanto ao desenvolvimento, implementação e uso.

Soluções e sistemas de IA podem operar de forma autônoma ou funcionar apenas como componentes de

[4] O titulo do artigo é: *A PROPOSTA REGULATÓRIA DA UNIÃO EUROPEIA PARA A INTELIGÊNCIA ARTIFICIAL (2ª. parte) – Sistemas de "risco inaceitável"*. Acessível em: https://jus.com.br/artigos/90817 .

produtos[5] colocados no mercado de consumo, aumentando os riscos à saúde e direitos fundamentais dos usuários. Por essa razão, os sistemas incluídos na categoria de "alto risco" (*high-risk*) se sujeitam a requisitos de conformidade e avaliação prévia de impacto, antes de serem colocados no mercado. É importante que os riscos de segurança que podem ser gerados por um produto, em razão de seus componentes digitais (incluindo sistemas de IA), sejam prevenidos e mitigados.

A classificação como um sistema de "alto risco" depende não somente da função que desempenha, mas fundamentalmente da finalidade para a qual foi concebido. À luz da "finalidade pretendida"[6], um sistema de IA, quer se trate de um produto autônomo ou componente de um produto, pode ser enquadrado nessa categoria se representar um alto risco de dano à saúde e segurança ou aos direitos fundamentais das pessoas, tendo em conta a gravidade do possível dano e probabilidade de sua ocorrência.

O Capítulo 1 do Título III do Regulamento traz um pequeno número de regras que possibilitam a identificação de sistemas de IA como de "alto risco". Basicamente, são identificadas duas categorias de sistemas de IA de "alto risco":

a) os que funcionam como **componentes de segurança** de produtos que estão sujeitos à avaliação de conformidade prévia por órgãos de controle;

b) outros sistemas independentes de IA com implicações em direitos fundamentais que estão explicitamente listados no Anexo III do Regulamento (art. 6°.).

---

[5] Um sistema de IA pode ser apenas parte de um produto ou ele mesmo ser o produto (*stand-alone AI systems*).

[6] "Finalidade pretendida" significa o uso para o qual um sistema de IA é pretendido pelo provedor, incluindo o contexto específico e as condições de uso, conforme as informações prestadas pelo fornecedor nas instruções de uso, materiais promocionais ou de vendas, declarações e documentação técnica.

Se um determinado produto, para ser colocado no mercado de consumo, está obrigado a se submeter previamente a um procedimento de avaliação de conformidade por algum órgão de controle governamental, segundo normas estabelecidas na legislação nacional para segurança de produtos, essa circunstância é suficiente para caracterizar o componente digital de segurança (sistema de IA) como de "alto risco". É o que ocorre, por exemplo, com certas máquinas, brinquedos, elevadores, equipamentos e sistemas de proteção destinados ao uso em atmosferas potencialmente explosivas, equipamentos de rádio, equipamentos de pressão, instalações por cabo, aparelhos que queimam combustíveis gasosos e equipamentos médicos. Todos os sistemas de IA, que funcionam como componentes de segurança desses produtos regulados, que se submetem ao controle de órgãos que realizam uma avaliação de conformidade, são considerados de alto risco. Se o próprio produto já tem que se submeter a uma avaliação prévia de conformidade com padrões de segurança, isso faz que o componente digital desse mesmo produto seja considerado como de "alto risco".

Além desses sistemas de IA que funcionam como **componentes de segurança** de produtos submetidos a avaliações prévias de conformidade, o Regulamento traz ainda uma lista de sistemas autônomos de IA (*stand-alone AI systems*) que são utilizados em várias áreas especificamente predefinidas. A lista está contida no Anexo III do Regulamento e se refere a sistemas de IA utilizados, dentre outras atividades e situações, nas seguintes: i) **infraestruturas críticas** que possam comprometer a vida ou integridade física das pessoas (p. ex., transportes); ii) **educação ou formação profissional**, que tenham o potencial de restringir o acesso à educação e à evolução profissional de alguém (p. ex., classificação de exames); iii) **componentes de segurança de produtos** (p. ex., cirurgia

assistida por robôs); iv) **emprego, gestão de trabalhadores e acesso ao trabalho por conta própria** (p. ex., análise de currículo em processos seletivos); v) **serviços públicos e privados essenciais** (p. ex., pontuação de crédito para concessão de empréstimos); vi) "**aplicação coercitiva da lei**" que possa interferir com os direitos fundamentais das pessoas (p. ex., avaliação da fiabilidade de provas); vii) **gestão da migração e do controle de fronteiras** (p. ex., verificação da autenticidade de documentos de viagem); e viii) **administração da justiça** e processos democráticos (p. ex., aplicação da lei em casos concretos) .

São classificados como de "alto risco" sistemas de IA empregados na gestão e operação de **infraestruturas críticas**, sobretudo quando destinados a serem utilizados como componentes de segurança na gestão e exploração da circulação rodoviária e no abastecimento de água, gás, aquecimento e eletricidade, uma vez que uma falha ou mau funcionamento pode colocar em risco a vida e a saúde das pessoas e levar a consideráveis interrupções na condução normal das atividades sociais e econômicas.

Também são considerados de "alto risco" os sistemas de IA usados na **educação** ou **formação profissional**, nomeadamente para determinar o acesso ou atribuir pessoas a instituições de ensino ou para avaliá-las em testes como pré-condição para sua educação, uma vez que podem determinar o curso educacional e profissional da vida de uma pessoa e, portanto, afetar sua capacidade de garantir seu sustento. Quando projetados e usados de maneira inadequada, tais sistemas podem violar o direito à educação e perpetuar padrões históricos de discriminação de certos grupos de pessoas.

Os sistemas de IA usados no **emprego, gestão de trabalhadores e acesso ao trabalho autônomo**, nomeadamente para o recrutamento e seleção de pessoas, para

a tomada de decisões sobre promoção e rescisão e para atribuição de tarefas, monitoramento ou avaliação de pessoas em relações contratuais relacionadas ao trabalho, também são classificados como de "alto risco", uma vez que esses sistemas podem ter um impacto significativo nas perspectivas de carreira futura e nos meios de subsistência dessas pessoas. Ao longo do processo de recrutamento e na avaliação, promoção ou retenção de pessoas em relações contratuais relacionadas ao trabalho, tais sistemas podem perpetuar padrões históricos de discriminação, por exemplo, contra mulheres, certos grupos etários, pessoas com deficiência ou pessoas de determinada raça, origem étnica ou orientação sexual. Os sistemas de IA usados para monitorar o desempenho e o comportamento das pessoas também podem afetar direitos relacionados à proteção de dados e privacidade individual, pois coletam vasta quantidade de dados pessoais.

Atualmente verifica-se uma rápida expansão do recrutamento de pessoas para postos de trabalho com base em ferramentas de inteligência artificial. Existem diversas modalidades de sistemas de IA e algoritmos voltados à seleção para o trabalho e emprego. A tarefa de seleção para vagas de emprego, que antes era realizada por recrutadores humanos, hoje é feita de modo automatizado, por meio de *softwares* de inteligência artificial, sobretudo nas grandes empresas.

Durante o processo seletivo, *softwares* de inteligência artificial avaliam vários aspectos da personalidade, capacidade e inteligência dos candidatos a emprego e decidem, sem interferência humana, se aprovam ou rejeitam os interessados aos postos oferecidos. Os próprios sistemas de IA montam as perguntas do processo de avaliação e, dependendo das respostas e da rapidez com que as perguntas são respondidas, situam os candidatos como

aprovados ou reprovados ao emprego. Atributos cognitivos e também emocionais são avaliados pelos algoritmos de inteligência artificial. Muitos dos programas possuem funcionalidades que gravam vídeos dos candidatos respondendo a perguntas durante entrevistas e, por esse meio, classificam certos aspectos do comportamento emocional do candidato.

As empresas que desenvolvem e comercializam esses algoritmos utilizados no processo de recrutamento para postos de trabalho garantem que testam seus sistemas contra vieses e discriminações, mas obviamente se sabe que os algoritmos não são indenes a falhas. Basta lembrar o caso que se tornou público em 2018, envolvendo o *software* de recrutamento utilizado pela Amazon, que incorporava preconceitos contra mulheres que disputavam os cargos. O sistema da Amazon ensinou a si mesmo que os candidatos do sexo masculino eram preferíveis porque costumavam ter mais experiência no setor de tecnologia. Como a maioria dos currículos enviados para a Amazon eram de homens interessados nas vagas - como acontece na maior parte da indústria tecnológica, onde predominam pessoas do sexo masculino -, o sistema entendeu que candidatos homens naturalmente eram mais aptos para os postos de trabalho na empresa. A simples menção ao sexo feminino era interpretada de forma negativa pelo algoritmo, que reduzia as chances das postulantes. A equipe responsável pelo desenvolvimento da ferramenta só percebeu o problema depois que já estava sendo usada[7].

[7] Ver notícia publicada no site *Olhar Digital*, em 10.10.18, acessível em: https://olhardigital.com.br/2018/10/10/noticias/inteligencia-artificial-da-amazon-exercitava-preconceito/

Os sistemas de IA usados para avaliar a **pontuação de crédito** (*credit scoring*)[8] ou **capacidade de crédito** de pessoas físicas são categorizados como de alto risco, uma vez que determinam o acesso a recursos financeiros ou serviços essenciais, como habitação, eletricidade e serviços de telecomunicações. Os sistemas de IA usados para fins de avaliação de solvência e classificação de crédito podem levar à discriminação de pessoas ou grupos e perpetuar padrões históricos de discriminação, com base em origens raciais ou étnicas, deficiências, idade, orientação sexual ou criar novas formas de impactos discriminatórios.

Outra área em que o uso de sistemas de IA merece qualificação como de "alto risco" é a relacionada com o acesso e gozo de certos **serviços e benefícios públicos**. As pessoas que solicitam ou recebem benefícios e serviços de assistência pública dependem normalmente desses benefícios e serviços e estão numa posição vulnerável em relação às autoridades públicas. Sistemas de IA utilizados para determinar se tais benefícios e serviços devem ser negados, reduzidos ou revogados, podem ter um impacto significativo na vida das pessoas e infringir seus direitos fundamentais, como o direito à proteção social, não-discriminação e dignidade humana. Sobretudo sistemas que são utilizados para estabelecer prioridades no atendimento a emergências devem ser classificados como de "alto risco",

---

[8] O ***Credit Score*** é uma técnica utilizada por bancos ou financeiras para medir os riscos na concessão de crédito, através da análise do histórico da pessoa que pede empréstimo. A *pontuação de crédito* pode indicar se o tomador de crédito pode ser um potencial inadimplente. Para analisar o *credit score*, os bancos ou financeiras utilizam uma base de dados com informações sobre o tomador de crédito, que envolvem rendimentos mensais, protestos, ações judiciais ou outras dívidas em aberto. Para facilitar a análise é possível recorrer a empresas privadas que trabalham fornecendo essas informações, a exemplo do Serviço de Proteção ao Crédito (SPC) e a Serasa Experian.

uma vez que tomam decisões em situações muito críticas para a vida e saúde de pessoas e seus bens.

A proposta de Regulamento também considera apropriado classificar como de "alto risco" sistemas de IA utilizados no contexto de **aplicação da lei**, onde a precisão, confiabilidade e transparência são particularmente importantes para evitar impactos adversos, manter a confiança pública e garantir responsabilidade e reparação efetiva (em caso de falha). Nessa categoria se incluem praticamente todos os sistemas utilizados por autoridades públicas para o fim de prevenção, detecção, investigação e persecução de delitos. O conjunto de sistemas utilizados por autoridades encarregadas da aplicação da lei apresentam elevados riscos para os direitos e liberdades fundamentais das pessoas, notadamente aqueles que desempenham as seguintes funções: a) avaliação de riscos individuais; b) polígrafos[9] e ferramentas semelhantes para detectar o estado emocional de uma pessoa; c) avaliação da confiabilidade das evidências (provas) em processos criminais; d) prevenir a ocorrência ou recorrência de um crime real ou potencial com base no perfil de pessoas físicas, ou avaliação de traços de personalidade e características ou comportamento criminoso passado de pessoas físicas ou grupos, para traçar o perfil no curso de detecção, investigação ou persecução de crimes. Esses tipos de sistemas são particularmente preocupantes porque podem proporcionar a vigilância, prisão ou privação de liberdade das pessoas, assim como produzir impactos adversos em garantias fundamentais.

Nessa área de investigação criminal e persecução criminal, se o sistema não for treinado com

[9] Um **polígrafo** ou detector de mentiras é um aparelho que mede e grava registros de diversas variáveis fisiológicas enquanto um interrogatório é realizado, supostamente para tentar identificar mentiras num depoimento.

dados confiáveis, não atender aos requisitos adequados em termos de precisão ou robustez, ou não for devidamente projetado e testado antes de ser colocado no mercado ou entrar em serviço, ele pode classificar as pessoas de forma discriminatória, incorreta ou injusta. Além disso, o exercício de importantes direitos processuais fundamentais, como o direito a um recurso efetivo e a um julgamento justo, bem como o direito de defesa e a presunção de inocência, pode ser dificultado, em particular quando tais sistemas de IA não são suficientemente transparentes, explicáveis e documentados.

O Regulamento proposto ainda considera como de "alto risco" sistemas utilizados em **serviços de imigração, asilo e controle de fronteiras**, porque alteram a situação e o estado de pessoas que em regra estão em posição particularmente vulnerável e que são dependentes do resultado das ações das autoridades públicas competentes. A precisão, natureza não discriminatória e transparência dos sistemas de IA usados nessas áreas são particularmente importantes para garantir o respeito dos direitos fundamentais das pessoas, nomeadamente direitos à livre circulação, não discriminação e proteção da vida privada.

Por fim, são também classificados como de "alto risco" sistemas relacionados com a **administração da Justiça**, em razão do impacto sobre direitos e liberdades individuais. Sistemas destinados a auxiliar autoridades judiciárias na pesquisa e interpretação de fatos e aplicação da lei, em razão dos riscos de possíveis vieses, erros e opacidade, são regulados por meio de exigências mais severas, a exemplo de todos os outros sistemas conceituados como de "alto risco". Alguns sistemas que são utilizados para realização de atividades administrativas auxiliares, no entanto, não são considerados como de "alto risco", porque não afetam decisões judiciárias em casos

específicos, como, p. ex., programas para anonimização ou armazenamento de dados, comunicação entre funcionários e outras tarefas meramente auxiliares.

O art. 7º. do Regulamento autoriza a *Comissão Europeia* a atualizar a lista do Anexo III, sempre que novos sistemas de IA coloquem em risco a segurança e saúde das pessoas ou tenham potencial de impacto negativo na órbita de seus direitos fundamentais.

**Requisitos obrigatórios para os sistemas de "alto risco"**

Para mitigar os riscos dos sistemas de IA colocados no mercado, certos requisitos obrigatórios são estipulados, tendo em conta a finalidade prevista para o sistema e de acordo com a gestão de risco adotada pelo provedor. Esses requisitos estão dispostos no Capítulo 3 do Título III. Alguns são estipulados também para usuários e outros participantes da cadeia de fornecimento dos sistemas (como, p. ex., importadores, distribuidores e representantes comerciais).

Para colocar no mercado consumidor da UE ou iniciar o funcionamento de um produto ou equipamento com algum componente ou programa de inteligência artificial (conceituado como de "alto risco"), o operador ou provedor deve implementar e manter um ***sistema de gerenciamento de risco*** (*risk management system*), que deve acompanhar e realizar testes de segurança durante todo o ciclo de vida do sistema de IA (art. 9º.) e mesmo antes de sua colocação no mercado. Ao usuário devem ser prestadas todas as informações relativas à existência de risco residual, não eliminado pelo sistema de gerenciamento de risco.

Ainda para mitigar os riscos à saúde, segurança e direitos fundamentais das pessoas, o Regulamento prevê outras exigências e condições para

colocação no mercado ou início de funcionamento de sistemas de IA de "alto risco". Os requisitos estão relacionados com a qualidade dos dados, a necessidade de documentação e manutenção de registros, transparência quanto ao funcionamento dos sistemas, dever de informação ao usuário, submissão à supervisão humana, robustez, acurácia e resistência dos sistemas a ciberataques.

A **qualidade dos dados** que alimentam um sistema de IA é fundamental para a sua adequada performance, especialmente quando são utilizados durante o processo de "treinamento", para evitar discriminações a certas categorias ou grupos de pessoas. Durante o processo de treinamento, os parâmetros de "apredizado da máquina" são determinados pelos dados utilizados nessa fase, daí a importância de serem livres de erros ou inexatidões. Os "dados de treinamento" são aqueles usados para treinar um sistema de IA por meio do ajuste de seus "parâmetros aprendíveis". Como se sabe, o *aprendizado automático* (ou *aprendizado de máquina*) envolve algoritmos que podem aprender com seus erros e fazer previsões com base em dados. Tais algoritmos operam construindo um modelo a partir de *inputs* amostrais a fim de fazer previsões ou decisões guiadas pelos dados, ao invés de simplesmente seguir inflexíveis e estáticas instruções programadas[10]. Essa característica denota a importância da qualidade dos dados para o processo de treinamento do sistema de IA.

O mesmo acontece durante as fases de validação e teste. Os conjuntos de dados de treinamento, validação[11] e teste[12] requerem a implementação de **governança**

---

[10] Enquanto que na *inteligência artificial* existem dois tipos de raciocínio (o indutivo, que extrai regras e padrões de grandes conjuntos de dados, e o dedutivo), o *aprendizado de máquina* só se preocupa com o indutivo. Cf. Wikipédia. Acessível em: https://pt.wikipedia.org/wiki/Aprendizado_de_m%C3%A1quina

[11] "Dados de validação" (*validation data*) são aqueles usados para fornecer uma avaliação da IA treinada e para ajustar seus parâmetros não aprendíveis e seu processo de aprendizagem. Os "dados de

**de dados** apropriada e práticas de gerenciamento (art. 10, 6). Devem ser completos e estar livres de erros, tendo em vista a finalidade pretendida do sistema (art. 10, 1 a 4).

A fim de proteger o direito de determinadas categorias ou grupos de pessoas contra vícios de discriminação, os provedores devem ser capazes de processar categorias especiais de dados pessoais, a fim de viabilizar o monitoramento, detecção e correção de vieses em sistemas de IA de alto risco (art. 10, 5).

Os desenvolvedores de sistemas de IA dentro do espaço europeu têm a vantagem de ter facilidade de acesso a conjuntos de dados de qualidade relacionados a seus respectivos campos de atividades. A União Europeia já regulamentou sua política de **governança de dados**, estabelecendo a criação de "espaços comuns de dados europeus", para facilitar o compartilhamento entre empresas e com o governo[13]. A criação desses "espaços comuns" será fundamental para fornecer informações confiáveis, responsáveis e não discriminatórias, viabilizando o acesso a dados de alta qualidade para o treinamento, validação e teste de sistemas de IA. Por exemplo, o "espaço europeu de dados de saúde" facilitará o treinamento de algoritmos de inteligência artificial nesses conjuntos de dados, de maneira a preservar a privacidade, segurança e transparência e garantir governança institucional adequada.

---

validação" podem ser um conjunto de dados separados ou parte do conjunto de dados de treinamento.

[12] "Dados de teste" são os usados para fornecer uma avaliação independente do sistema de IA já treinado e validado, a fim de confirmar o desempenho esperado desse sistema antes de sua colocação no mercado ou entrada em serviço.

[13] Para saber mais sobre a política de governança de dados da União Europeia, sugerimos a leitura de nosso artigo "O GOVERNANCE DATA ACT – A estratégia europeia para manter a soberania sobre os dados e conter o domínio dos mercados digitais pelas "Big Techs", publicado na Revista Jus Navigandi, em 31.12.20. Disponível em: https://jus.com.br/artigos/87649

Outro requisito estabelecido no Regulamento, em relação ao desenvolvimento e operação de sistemas de IA de "alto risco", é a necessidade de **documentação** de todo o ciclo de vida do projeto. Ter informações sobre como os sistemas de IA de alto risco foram desenvolvidos e como eles executam as tarefas é essencial para verificar a conformidade com as regras estabelecidas no Regulamento. A documentação técnica deve conter as informações necessárias para avaliar a conformidade do sistema de IA com as exigências regulamentares. Essas informações devem incluir as características gerais, capacidades e limitações do sistema, algoritmos, dados, processos de treinamento, teste e validação usados, bem como documentação sobre o sistema de gestão de risco relevante. A documentação técnica deve ser elaborada e estar disponível antes do sistema ser colocado no mercado e deve ser mantida atualizada (art. 11, 1 a 3).

A **manutenção de registros** sobre o funcionamento e operação dos sistemas de IA de alto risco também é um requisito estipulado no Regulamento. Os sistemas de IA de alto risco devem ser projetados e desenvolvidos com recursos que permitam o registro automático de eventos ("logs") enquanto estão operando. As capacidades de registro devem garantir um nível de rastreabilidade do funcionamento do sistema de IA ao longo do seu ciclo de vida (art. 12, 1 a 3). Para certas categorias de sistemas de IA de alto risco, o Regulamento exige que as capacidades de registro devem fornecer, no mínimo: (a) registro do período de cada uso do sistema (data e hora de início e término de cada uso); (b) o banco de dados de referência contra o qual os dados de entrada foram verificados pelo sistema; (c) os dados de entrada para os quais a pesquisa resultou em uma correspondência;

d) a identificação das pessoas físicas envolvidas na verificação dos resultados (art. 12, 4).

Para lidar com a opacidade que pode tornar certos programas e algoritmos incompreensíveis ou complexos para pessoas físicas, um certo grau de **transparência** é exigido para sistemas de IA de alto risco. Devem ser projetados e desenvolvidos de forma a garantir que a operação seja suficientemente transparente para permitir que os usuários interpretem a saída do sistema e possam usá-la adequadamente (art. 13, 1). Os sistemas de IA de alto risco devem, portanto, ser acompanhados por documentação e instruções de uso e incluir informações concisas e claras, inclusive em relação a possíveis riscos aos direitos fundamentais e vícios de discriminação, onde apropriado (art. 13, 2). As informações que devem ser prestadas aos usuários, devem especificar:

(a) a identidade e os detalhes de contato do provedor e, quando aplicável, de seus representantes autorizados;

(b) as características, capacidades e limitações de desempenho do equipamento com sistema de IA de alto risco, incluindo:

(i) a finalidade pretendida;

(ii) o nível de precisão, robustez e cibersegurança contra os quais o sistema de IA de alto risco foi testado e validado e que pode ser esperado, bem como qualquer circunstância conhecida e previsível que pode ter impacto sobre o nível esperado de precisão, robustez e segurança cibernética;

(iii) qualquer circunstância conhecida ou previsível, relacionada ao uso do sistema de IA de alto risco, que, de acordo a finalidade pretendida ou sob condições de mau uso, pode levar a riscos para a saúde e segurança ou direitos fundamentais das pessoas;

(iv) seu desempenho em relação às pessoas ou grupos de pessoas para as quais o sistema se destina a ser usado;

(v) quando apropriado, especificações sobre os dados de entrada (*input data*) ou qualquer outra informação relevante em termos de conjuntos de dados de treinamento, validação e teste utilizados, levando em consideração a finalidade pretendida do sistema de IA.

(c) as mudanças no sistema de IA de alto risco ou no seu desempenho que foram pré-determinadas pelo provedor no momento da avaliação inicial de conformidade, se houver;

(d) as medidas de supervisão humana referidas no Artigo 14, incluindo as medidas técnicas postas em prática para facilitar a interpretação dos resultados do sistema de IA pelos usuários;

(e) a vida útil esperada do sistema de IA de alto risco e todas as medidas de cuidado e manutenção para garantir o adequado funcionamento, inclusive no que diz respeito a atualizações de *software*.

Como fica claro, nesse ponto o Regulamento estabelece um "direito à informação", em favor do usuário de um sistema de IA. A normatização garante direitos básicos de informação ao usuário, concretizando o *princípio da transparência informativa*. Esse direito à informação envolve os parâmetros, regras e instruções nos quais se baseiam os algoritmos ou sistemas de inteligência artificial. Trata-se de conjunto de preceitos de significado inovador, que se aplica como dever do provedor de qualquer sistema de inteligência artificial de "alto risco". O reconhecimento desse "direito à informação" transfere para o provedor a obrigação de informar ao usuário aspectos relevantes do funcionamento do sistema de IA. O elemento decisivo e original proporcionado por esse

"direito à informação" é a transparência de certos parâmetros e critérios do funcionamento do sistema de IA. Isso proporciona questionar a neutralidade desses parâmetros como discriminatórios ou impeditivos do exercício de direitos fundamentais. O "direito à informação" sobre o algoritmo viabiliza a eliminação de elementos discriminatórios e obstáculos ao exercício de direitos fundamentais.

Para minimizar os riscos à saúde, segurança e direitos fundamentais de usuários e terceiros, o Regulamento também exige que os sistemas de IA de "alto risco" sejam projetados e desenvolvidos de maneira apropriada à **supervisão humana**. Para tanto, os sistemas devem incorporar ferramentas de interface homem-máquina apropriadas, que permitam a supervisão por uma pessoa humana durante o período de uso do sistema. O sistema deve ser desenhado de forma a proporcionar supervisão humana desde sua concepção e antes de ser colocado no mercado ou iniciar seu funcionamento. As ferramentas para supervisão têm que ser construídas e embutidas dentro do sistema de maneira que a pessoa física responsável pela tarefa de supervisão seja capaz de, ao identificar qualquer indício de anomalias ou disfunções, intervir na operação ou suspender o funcionamento (art. 14, itens 1 a 5).

Os sistemas de IA de alto risco devem obedecer a um nível apropriado de **precisão**, **robustez** e **segurança cibernética** e funcionar com esse mesmo padrão ao longo de todo ciclo de vida. O nível e métricas de precisão devem ser comunicados aos usuários. A exigência de nível apropriado de segurança cibernética atribui ao provedor o dever de oferecer não apenas garantia de funcionamento isento de falhas, erros e inconsistências, mas também proteção contra ações maliciosas de terceiros que possam

comprometer a regular performance do sistema de IA de alto risco (art. 15, itens 1 a 4).

Recife, 21 de maio de 2021.