



**OS FUNDAMENTOS DOS MODELOS DE LINGUAGEM DA INTELIGÊNCIA
ARTIFICIAL: A ÉTICA COMO PRINCÍPIO E DIRETRIZ PARA O USO DAS IAS
NA ADMINISTRAÇÃO PÚBLICA**

**THE FUNDAMENTALS OF ARTIFICIAL INTELLIGENCE LANGUAGE MODELS:
ETHICS AS A PRINCIPLE AND GUIDELINE FOR THE USE OF AIs IN PUBLIC
ADMINISTRATION**

Luis Henrique Duarte¹

Camille Lima Reis²

Ana Carolina Dantas Rocha³

Filipe Lôbo Gomes⁴

RESUMO: A adoção da Inteligência Artificial (IA) na Administração Pública, como nos sistemas “Poupinha” e “Alice”, visa eficiência e melhor atendimento ao cidadão. No entanto, exige cuidados éticos para evitar violações de direitos, discriminações e ampliar desigualdades. Desafios como vieses, falta de transparência e proteção de dados demandam governança robusta. Inspirado em Aristóteles, Kant e na ética algorítmica, o estudo propõe uso ético, transparente e responsável da IA. Contrasta-se o modelo regulatório europeu (AI Act) com o cenário brasileiro, sugerindo a criação de estruturas permanentes de controle ético e governança interdisciplinar para proteger os princípios constitucionais.

PALAVRAS-CHAVE: inteligência artificial; modelos de linguagem; administração pública; ética algorítmica; direito administrativo.

ABSTRACT: The adoption of Artificial Intelligence (AI) in Public Administration, as seen in systems like “Poupinha” and “Alice,” aims to enhance efficiency and improve citizen services. However, it requires ethical safeguards to prevent rights violations, discrimination, and increased inequalities. Challenges such as algorithmic bias, lack of transparency, and data

¹Formado pelo Centro Universitário de Maceió (UNIMA). Acadêmico de Direito pelo Centro Universitário CESMAC. E-mail: henriquedvc@hotmail.com.

²Mestre em direito público pela Universidade Federal de Alagoas (UFAL). Professora e pesquisadora. Advogada especialista em direito médico e da saúde. E-mail: millelreis@gmail.com.

³Graduação em Engenharia de Controle e Automação pelo Centro Universitário de Maceió (UNIMA). Acadêmica de Direito pelo Centro Universitário CESMAC. E-mail: carolinadantars24@gmail.com.

⁴Pós-doutorando em Direito pela UFPE e pela Faculdade de Direito da Universidade de Coimbra. Doutor em Direito pela UFPE. Professor no mestrado e na graduação em Direito da UFAL e do CESMAC. Professor da Escola Superior da Magistratura de Alagoas. Pesquisador líder do grupo de pesquisa Núcleo de Estudos da Regulação e da Nova Gestão Pública. Sócio fundador do Instituto de Direito Administrativo do Estado de Alagoas - IDAA. Vice-Diretor da FDA-UFAL. Procurador-Geral do TJ-AL. E-mail: filipelobo@uol.com.br.

protection demand strong governance. Inspired by Aristotle, Kant, and algorithmic ethics, the study advocates for ethical, transparent, and responsible use of AI. It contrasts the European regulatory model (AI Act) with the Brazilian context, suggesting the creation of permanent ethical oversight structures and interdisciplinary governance to uphold constitutional principles.

KEYWORDS: artificial intelligence; language models; public administration; algorithmic ethics; administrative law.

1 INTRODUÇÃO

A integração da inteligência artificial (IA), especialmente dos modelos de linguagem, à Administração Pública representa uma das transformações mais significativas do século XXI, trazendo oportunidades e desafios para o Direito Administrativo. A capacidade desses modelos de processar e gerar linguagem natural tem potencial para otimizar processos administrativos, aumentar a eficiência e melhorar o atendimento ao cidadão, como demonstrado por iniciativas como o chatbot “Poupinha” do INSS e o sistema “Alice” do Tribunal de Contas da União (TCU). Contudo, a adoção dessas tecnologias exige um equilíbrio delicado para assegurar a conformidade com os princípios constitucionais da Administração Pública, como legalidade, impensoalidade, moralidade, publicidade e eficiência, previstos no artigo 37 da Constituição Federal do Brasil.

O problema de pesquisa que guia este estudo é: Como integrar modelos de linguagem de inteligência artificial na administração pública brasileira de forma ética e responsável, garantindo a conformidade com os princípios constitucionais e evitando riscos como vieses algorítmicos e opacidade decisória? A pesquisa parte da premissa de que a IA, embora traga benefícios como agilidade e eficiência, também pode ampliar desigualdades, violar direitos e gerar discriminações se não for implementada com padrões éticos robustos. Assim, o estudo busca não apenas compreender os conceitos técnicos e jurídicos envolvidos, mas também propor soluções práticas para harmonizar modernidade tecnológica com os postulados clássicos do Direito Administrativo, com análise crítica dos riscos e implicações. Os modelos de linguagem, baseados em algoritmos complexos de aprendizado de máquina, como a arquitetura Transformer, oferecem soluções inovadoras, mas também levantam questões éticas e jurídicas cruciais. Problemas como vieses algorítmicos, opacidade das decisões (a chamada “caixa preta”) e a proteção de dados pessoais, regulada pela Lei Geral de Proteção de Dados (LGPD), demandam uma governança robusta e interdisciplinar. Além

disso, a reflexão ética, inspirada em pensadores como Aristóteles e Kant, e adaptada ao contexto da ética algorítmica, é essencial para garantir que a IA seja utilizada de forma transparente, responsável e alinhada ao interesse público.

Este trabalho busca explorar os fundamentos técnicos, éticos e jurídicos dos modelos de linguagem, analisando suas aplicações práticas na Administração Pública e os desafios que impõem ao ordenamento jurídico brasileiro. Por meio de uma abordagem interdisciplinar, que conecta o Direito, a tecnologia e a ética, o estudo propõe diretrizes para uma implementação responsável da IA, visando fortalecer a eficiência administrativa sem comprometer os valores democráticos e os direitos fundamentais.

Esta pesquisa adota uma abordagem qualitativa e interdisciplinar, combinando elementos do Direito Administrativo, da Ética e da Ciência da Computação. Utiliza-se método bibliográfico e documental, com análise de doutrina, legislação (como a Constituição Federal, LGPD e AI Act da UE) e casos práticos (ex.: "Poupinha" e "Alice"). A delimitação foca no contexto brasileiro, com comparação ao modelo europeu para destacar lacunas regulatórias. A análise crítica é realizada por meio de exame dedutivo, partindo de princípios gerais (éticos e constitucionais) para aplicações específicas na administração pública. Não se empregam métodos quantitativos, priorizando-se a interpretação normativa e propositiva de soluções.

2 OS FUNDAMENTOS DOS MODELOS DE LINGUAGEM DA INTELIGÊNCIA ARTIFICIAL E A ADMINISTRAÇÃO PÚBLICA

2.1 MODELOS DE LINGUAGEM DA INTELIGÊNCIA ARTIFICIAL: UM MAPA CONCEITUAL NECESSÁRIO

Um dos grandes desafios do Direito Administrativo contemporâneo está na incorporação de novas tecnologias ao aparato estatal, sem comprometer os princípios que regem a Administração Pública, como legalidade, impessoalidade, moralidade, publicidade e eficiência, previstos no artigo 37 da Constituição Federal do Brasil. A ascensão da inteligência artificial (IA), especialmente em sua vertente mais sofisticada, os modelos de linguagem, intensifica esse desafio, exigindo do Direito uma atualização normativa que evite a obsolescência e uma reflexão profunda sobre os impactos éticos, sociais e jurídicos dessas inovações. No Brasil, iniciativas como o chatbot "Poupinha" do INSS, que auxilia cidadãos em questões previdenciárias, e o uso de IA pelo Tribunal de Contas da União (TCU) para análise de licitações demonstram o potencial da tecnologia para aumentar a eficiência. Contudo, casos internacionais, como o algoritmo COMPAS nos Estados Unidos, que

apresentou vieses raciais em decisões judiciais, alertam para os riscos de uma implementação sem critérios éticos e jurídicos, especialmente em relação à impessoalidade e à igualdade.

2.1.1 CONCEITO DE INTELIGÊNCIA ARTIFICIAL

Inteligência Artificial (IA) refere-se ao desenvolvimento de sistemas computacionais capazes de realizar tarefas que normalmente exigem inteligência humana, como aprendizado, resolução de problemas, tomada de decisão e percepção. Segundo Russell e Norvig (2021), a IA pode ser dividida em categorias como IA estreita (focada em tarefas específicas, como modelos de linguagem) e IA geral (capaz de realizar qualquer tarefa intelectual humana, ainda em desenvolvimento). No contexto da Administração Pública, a IA estreita é predominante, com modelos de linguagem desempenhando papéis cruciais em automação de processos e interação com cidadãos. Esses sistemas utilizam aprendizado de máquina, redes neurais e processamento de linguagem natural (PLN) para interpretar e gerar textos, oferecendo soluções escaláveis para desafios administrativos.

Antes de explorar a aplicação desses modelos na Administração Pública, é essencial compreender o que são os modelos de linguagem e por que seus fundamentos técnicos e éticos devem ser analisados sob a ótica do Direito. As diretrizes da União Europeia destacam princípios como transparência, responsabilidade, segurança, privacidade, proteção de dados e não discriminação como fundamentais para o uso ético da IA. Esses princípios são cruciais para garantir que a tecnologia sirva ao interesse público sem violar direitos fundamentais, como o direito à igualdade previsto no artigo 5º da Constituição Brasileira.

Os modelos de linguagem são estruturas computacionais baseadas em IA que processam, geram e compreendem a linguagem natural, ou seja, a forma como seres humanos se comunicam por meio de palavras e textos. Para o jurista, podem ser comparados a um "assessor jurídico virtual" que lê, interpreta e redige documentos, responde perguntas ou traduz textos automaticamente. Conforme Jurafsky e Martin (2023), esses modelos são essenciais para tarefas como tradução automática, geração de textos (e.g., relatórios administrativos) e análise sintática (e.g., identificação de padrões em decisões judiciais). Eles são amplamente utilizados em assistentes virtuais, como a Alexa, e em sistemas de busca, como o Google.

Esses modelos funcionam como "cérebros artificiais" que aprendem a prever sequências de palavras com base em grandes quantidades de dados textuais, chamados corpora. Pode-se compará-los a um jurista que, após estudar milhares de acórdãos, prevê o

argumento mais adequado para um caso com base em padrões identificados. Essa capacidade é alcançada por meio de algoritmos de aprendizado de máquina, que ajustam bilhões de parâmetros (valores numéricos que definem o comportamento do modelo). Python é a principal linguagem de programação usada para criar e aplicar esses modelos, devido à sua sintaxe clara, que se assemelha à linguagem natural, e à sua vasta biblioteca de ferramentas, como NLTK, spaCy, Transformers (Hugging Face) e TensorFlow.

Segundo o Cass do IEEE Spectrum (2025), Python lidera o ranking das linguagens de programação mais usadas no mundo, impulsionado por seu papel em machine learning, data science e processamento de linguagem natural (PLN). Para o jurista, Python é como um "livro de leis" acessível, que permite a profissionais sem formação técnica desenvolver soluções de IA. Por exemplo, um procurador pode usar a biblioteca spaCy para analisar automaticamente petições administrativas, identificando termos jurídicos relevantes.

O avanço de modelos de linguagem de larga escala, como o GPT (Generative Pre-trained Transformer) da OpenAI, revolucionou o PLN. Esses modelos, com bilhões de parâmetros, realizam tarefas complexas, como redigir pareceres jurídicos ou responder a consultas de cidadãos. A arquitetura Transformer, introduzida por Vaswani *et al.* (2017), é o núcleo desses modelos, funcionando como uma rede neural que processa textos de forma paralela, capturando relações complexas entre palavras. Para o Direito, essa arquitetura é relevante porque determina a precisão das respostas, mas também apresenta o desafio da "caixa preta", ou seja, a dificuldade de explicar como o modelo chegou a uma decisão, o que pode conflitar com o princípio da motivação previsto na Lei nº 9.784/1999.

A acessibilidade de Python democratiza a IA. Iniciativas como o curso "Natural Language Processing with Python" da Universidade de Stanford (Manning, 2022) e a documentação aberta de ferramentas como spaCy permitem que juristas e gestores públicos desenvolvam aplicações de IA. Plataformas como o Hugging Face oferecem modelos pré-treinados, úteis para tarefas como análise de petições ou tradução de documentos oficiais.

2.1.2 LIMITAÇÕES TÉCNICAS E IMPLICAÇÕES JURÍDICAS DOS MODELOS DE LINGUAGEM

Os modelos de linguagem oferecem benefícios, mas suas limitações técnicas geram desafios jurídicos para o Direito Administrativo. Primeiro, esses modelos dependem de grandes corporações que podem conter vieses sociais, como discriminações de gênero ou raça. Bender *et al.* (2021) alertam que modelos como o GPT-3 podem perpetuar estereótipos se os

dados não forem auditados. Na Administração Pública, isso é crítico: um modelo enviesado poderia discriminar minorias em processos seletivos automatizados, violando o princípio da imparcialidade.

A opacidade dos modelos, conhecida como "caixa preta", é outro desafio. Como as decisões são baseadas em cálculos complexos, é difícil explicar por que o modelo chegou a uma resposta, o que conflita com o princípio da motivação (Lei nº 9.784/1999, art. 50). Por exemplo, se um modelo rejeita um pedido de benefício previdenciário, a falta de explication clara pode invalidar a decisão administrativamente.

O uso de dados pessoais também levanta questões de conformidade com a Lei Geral de Proteção de Dados (LGPD – Lei nº 13.709/2018). Modelos que processam informações sensíveis devem respeitar os princípios de finalidade, necessidade e segurança. A União Europeia, com o Regulamento Geral de Proteção de Dados (GDPR), exige que sistemas de IA sigam o conceito de privacy by design. Além disso, o alto custo energético para treinar modelos levanta questões de sustentabilidade, exigindo que órgãos públicos respeitem o princípio da eficiência.

2.2 OS FUNDAMENTOS ÉTICOS E JURÍDICOS DOS MODELOS DE LINGUAGEM

A complexidade da IA, com sua capacidade de aprendizado e autonomia, impõe novos contornos ao controle da atuação administrativa, exigindo critérios técnicos e jurídicos que garantam o uso legítimo, transparente e humanizado dessas ferramentas. A implementação de modelos de linguagem na Administração Pública requer uma reflexão profunda sobre seus fundamentos éticos e jurídicos, como transparência, responsabilização e respeito aos direitos fundamentais.

Essa reflexão constroi a ponte entre inovação tecnológica e interesse público, garantindo compatibilidade com os valores democráticos. Callejón (2020) defende a necessidade de "constitucionalizar o algoritmo", submetendo a IA aos princípios constitucionais, enquanto se adapta a Constituição ao mundo digital. No Brasil, isso implica alinhar o uso de IA aos princípios do artigo 37, especialmente legalidade e publicidade.

A regulamentação da IA exige marcos regulatórios flexíveis e rigorosos. O AI Act da União Europeia (2021) classifica sistemas de IA por níveis de risco, com exigências mais rigorosas para aplicações administrativas. A seguir, detalha-se a classificação de sistemas de IA por nível de risco conforme o AI Act (UE 2024/1689):

- a) Risco Inaceitável (vermelho): Sistemas que apresentam ameaças graves aos direitos

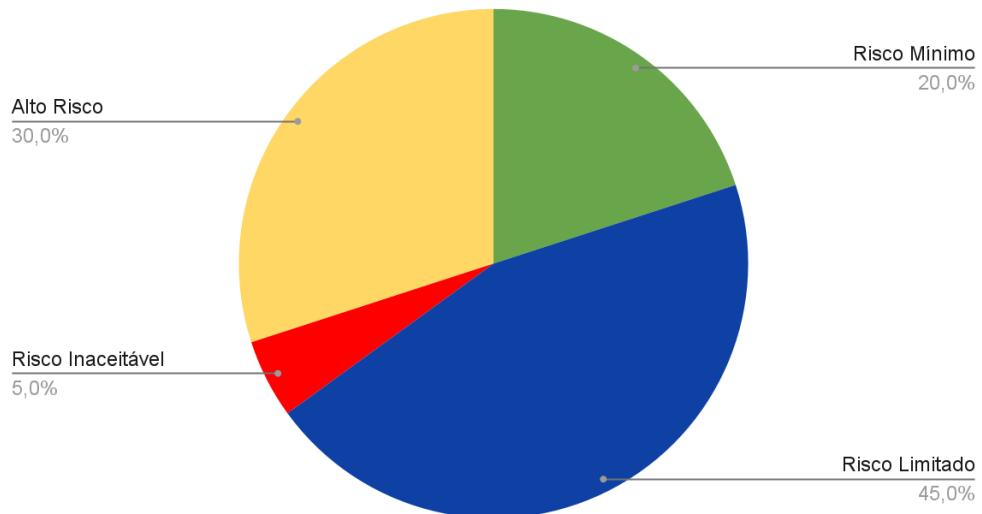
fundamentais, como manipulação comportamental em larga escala ou vigilância indiscriminada. Esses sistemas são proibidos, exceto em casos específicos, como segurança nacional.

- b) Alto Risco (amarelo): Sistemas que impactam significativamente a segurança ou os direitos fundamentais, como IA usada em decisões judiciais, licitações ou gestão de benefícios sociais. Exigem auditorias rigorosas, transparência e conformidade com padrões éticos.
- c) Risco Limitado (azul): Sistemas que requerem transparência mínima, como chatbots de atendimento ao cliente. Devem informar os usuários que estão interagindo com IA.
- d) Risco Mínimo (verde): Sistemas com impacto insignificante, como ferramentas de recomendação de conteúdo. Não possuem exigências regulatórias específicas.

O gráfico a seguir ilustra essa classificação com base no regulamento europeu:

Gráfico 1 – Classificação de Sistemas de IA por Nível de Risco (AI Act UE 2024/1689)

Classificação de Sistemas de IA por Nível de Risco



Fonte: Elaboração pelo autor (2025)

O impacto desta classificação reside em sua capacidade de destacar a proporção e a gravidade dos diferentes níveis de risco associados aos sistemas de inteligência artificial, fornecendo uma base visual para a formulação de políticas públicas. A predominância do Risco Limitado sugere que a maioria das aplicações de IA pode ser integrada à Administração Pública com regulamentações moderadas, promovendo eficiência sem comprometer a segurança. Por outro lado, a presença de categorias como Risco Inaceitável e Alto Risco alerta para a necessidade de controles rigorosos, especialmente em processos sensíveis, como

decisões judiciais ou licitações, evitando violações de princípios como impessoalidade e igualdade. Esse gráfico, portanto, serve como uma ferramenta essencial para orientar a governança da IA, incentivando a adoção de medidas preventivas e a alocação de recursos para auditorias e treinamentos éticos.

No Brasil, o Projeto de Lei nº 2.338/2023 enfatiza transparência e proteção de direitos. Contudo, a rápida evolução tecnológica exige governança adaptativa. Floridi *et al.* (2020) sugerem princípios éticos, como beneficência e justiça, além de auditoria contínua e participação pública.

2.2.1 GOVERNANÇA INTERDISCIPLINAR DA IA NO CONTEXTO BRASILEIRO

No Brasil, a governança da IA na Administração Pública deve ser interdisciplinar, envolvendo juristas, cientistas de dados e sociedade civil. O TCU, por exemplo, criou a plataforma "Alice" para análise de dados em auditorias, mas sua implementação exige comissões que monitorem vieses e garantam transparência. A participação pública, como em consultas abertas, é essencial para alinhar a IA ao interesse público, respeitando o princípio da publicidade. Ferramentas como o MLflow, que rastreia experimentos de IA, podem facilitar auditorias, garantindo conformidade com a LGPD.

2.3 APLICAÇÕES PRÁTICAS NA ADMINISTRAÇÃO PÚBLICA E DESAFIOS JURÍDICOS

Os modelos de linguagem transformam a Administração Pública ao otimizar processos. O "Poupinha" do INSS e o sistema "Alice" do TCU são exemplos brasileiros, enquanto Singapura usa IA para análise de políticas públicas. Esses sistemas, porém, desafiam o princípio da motivação, exigindo justificativas claras para decisões automatizadas. A conformidade com a LGPD é outro desafio, especialmente na proteção de dados sensíveis. Ferramentas como o MLflow podem ajudar, mas é necessário um marco regulatório robusto para garantir transparência e prestação de contas.

2.3.1 DESAFIOS DE IMPLEMENTAÇÃO E INCLUSÃO DIGITAL NA ADMINISTRAÇÃO PÚBLICA BRASILEIRA

A integração de IA em serviços como atendimento ao cidadão e gestão de recursos

humanos pode reduzir custos operacionais, mas exige a capacitação de servidores para operar e auditar esses sistemas. A experiência da Estônia, que utiliza plataformas digitais baseadas em IA para serviços públicos, oferece um modelo a ser estudado, destacando a importância de infraestrutura tecnológica e alfabetização digital da população. No entanto, no Brasil, desafios como desigualdade de acesso à internet e resistência cultural à automação devem ser superados para maximizar os benefícios da IA, reforçando a necessidade de políticas inclusivas e investimentos em conectividade.

Assim, o uso de IA na Administração Pública brasileira exige não apenas marcos normativos robustos, mas também a criação de estruturas institucionais permanentes de controle e auditoria ética, para garantir a compatibilidade entre inovação tecnológica e os direitos fundamentais consagrados na Constituição.

3 ÉTICA: PRINCÍPIO E DIRETRIZ PARA CANALIZAR À IA E A APU

3.1 ÉTICA ANTIGA: REFLEXÕES DE SÓCRATES E KANT

A distinção entre ética e moral tem gerado debates intensos na história da filosofia. Embora frequentemente tratadas como sinônimos, essas categorias conceituais possuem nuances relevantes. Enquanto a ética se refere a uma reflexão individual sobre a ação justa e boa, a moral está mais ligada às normas coletivas estabelecidas socialmente. Nesse contexto, ética e moral não são opostas, mas interdependentes, cada uma com suas especificidades. Conforme Santos (2021), a ética, como ressalta Aristóteles, orienta o indivíduo à busca da excelência e da virtude como condições para uma vida boa e feliz. Candioto (2023) destaca que Aristóteles foi pioneiro ao classificar a ética como uma disciplina filosófica prática, diferenciando-a dos saberes teóricos. Sua concepção propõe que a ética sirva de guia para ações moderadas com vistas ao bem-viver. Para o filósofo, todo conhecimento visa a um fim, e esse fim é o bem; assim, todas as ações humanas tendem naturalmente à realização do bem.

A ética aristotélica é, portanto, teleológica, isto é, fundamentada na ideia de finalidade. Em *Ética a Nicômaco*, Aristóteles defende que a virtude é uma disposição adquirida pela prática e tem como meta a felicidade. Tal pensamento evidencia que a ética é essencialmente uma busca racional e contínua pela excelência no agir, conforme Barbosa (2020).

Já na modernidade, Ogar e Bassey (2023) apontam que Immanuel Kant

introduz uma nova abordagem ao propor uma ética deontológica, fundamentada no dever e na razão. Na Fundamentação da Metafísica dos Costumes (1785), Kant estabelece o princípio moral supremo como sendo válido independentemente das consequências. A ação ética, portanto, não depende de inclinações naturais ou do prazer, mas da conformidade com a lei moral revelada pela razão. Fajardo (2021) complementa que Kant rejeita o empirismo ético ao afirmar que os conceitos morais não podem ser derivados da experiência, pois esta não possui universalidade nem necessidade. Assim, apenas princípios a priori, acessíveis pela razão, podem sustentar uma ética verdadeiramente objetiva. Nesse sentido, a vontade humana é livre porque pode agir segundo a razão e não por meras inclinações.

A discussão sobre a moralidade, ao longo da história da filosofia, tem sido intrinsecamente ligada ao conceito de sujeito moral. A determinação do que constitui um sujeito moral é fundamental para a ética, pois dela derivam as noções de direitos e responsabilidades. Caliendo (2021) destaca que São Boécio (480-524), por exemplo, revolucionou a filosofia ao definir a pessoa como uma 'substância individual de natureza racional'. Essa formulação, que enfatiza a singularidade e a racionalidade, foi um pilar para o surgimento do conceito moderno de indivíduo e da dignidade da pessoa humana. O estudioso também aponta que Boécio destacou que a dignidade humana não se fundamenta apenas na dimensão racional, mas na capacidade de fazer deliberações sobre o próprio agir, escolhendo o bem ou o mal.

Posteriormente, Caliendo (2021) revela que Ricardo de São Vitor (séc. XII) aprofundou essa noção ao caracterizar a pessoa como uma "existência incomunicável da natureza divina" (*persona est divinae naturae incommunicabilis existentia*), ressaltando a unicidade ontológica e a singularidade não compartilhável de cada indivíduo. O autor acrescenta que São Tomás de Aquino (1225-1274) adicionou a liberdade como elemento fundamental, definindo a pessoa como um ente especial que possui "domínio de seus atos e não são apenas movidas na ação como as outras, mas agem por si mesmas".

Nesse percurso, Caliendo (2021) argumenta que Kant consolidou a autonomia como o conceito central da moralidade. Para o filósofo, a autonomia é a "propriedade que a vontade possui de ser lei para si mesma (independentemente da natureza dos objetos do querer)". A vontade pura, livre de coações externas e inclinações empíricas, é a fonte da ação moral, conferindo dignidade ao ser racional. Caliendo (2021) explica que a moralidade, em Kant, é a relação das ações com a autonomia da vontade, ou seja, com a capacidade de o ser racional "determinar para si mesmo a lei de seu agir". Este princípio universal implica que a humanidade e toda natureza racional são fins em si mesmos, e não meros meios.

Embora o formalismo kantiano tenha sido alvo de críticas, como as de Hegel, que apontaram para a ausência de um conteúdo concreto ao dever, e as de Anscombe e MacIntyre, que questionaram a recusa em fundamentar a moralidade na natureza humana e a negligência das "virtudes do reconhecimento da dependência", Caliendo (2021) observa que a centralidade da autonomia para a moralidade permaneceu firmemente estabelecida.

A partir dessa base filosófica, ao avançarmos para a ética algorítmica, o desafio premente, segundo Caliendo (2021), é investigar se a Inteligência Artificial pode, de fato, ascender ao status de sujeito moral. Para que a IA seja considerada um "agente moral" autêntico, ela precisaria satisfazer esses critérios filosóficos clássicos, demonstrando racionalidade, autonomia na tomada de decisões, intencionalidade (agir por vontade própria e não por mera programação), e possivelmente a capacidade de sentir emoções e de assumir responsabilidade por seus atos. O autor conclui que essa transição de um mero instrumento para um potencial sujeito moral redefine as diretrizes éticas e jurídicas para o uso das IAs na Administração Pública.

No cenário contemporâneo, a ascensão da Inteligência Artificial (IA) tem imposto novas e complexas questões ao campo da ética, ampliando a necessidade de uma reflexão aprofundada que transcende a mera aplicação de princípios morais a sistemas tecnológicos. Caliendo explica que, tradicionalmente, a "ética artificial" pode ser compreendida em diferentes níveis: a ética aplicada à IA, que estabelece limites e diretrizes para sua pesquisa e desenvolvimento, e a ética decorrente da aplicação da IA, que analisa as consequências de seu uso. No entanto, o estudioso enfatiza que este artigo busca ir além, aprofundando-se na terceira e mais desafiadora acepção: a possibilidade filosófica de uma ética intrínseca à própria Inteligência Artificial, que emerja de sua vontade e racionalidade, transformando-a em um sujeito moral autêntico.

3.2 ÉTICA ALGORÍTMICA: PRINCÍPIOS E DESAFIOS NA ERA DA INTELIGÊNCIA ARTIFICIAL

O avanço da transformação digital tem impulsionado o uso de técnicas baseadas em Inteligência Artificial (IA) por instituições públicas e privadas, incorporando essas tecnologias em tarefas cotidianas sem que, muitas vezes, o usuário perceba. Desde compras digitais até recomendações em plataformas de streaming, a IA tornou-se onipresente. Durante a pandemia de COVID-19, esse processo foi ainda mais acelerado. No entanto, apesar dos ganhos operacionais promovidos especialmente pelo aprendizado de máquina, especialistas

têm apontado preocupações com os vieses e preconceitos reproduzidos por algoritmos inteligentes, o que evidencia a necessidade de atenção ética no desenvolvimento e aplicação desses sistemas (Garcia, 2020).

Inglada-Galiana, Corral-Gudiño e Miramontes-González (2024) explica que a presença cada vez mais significativa da IA nas esferas humanas não apenas reformula os aspectos tecnológicos, mas também desafia os fundamentos morais e éticos sobre os quais organizamos nossa sociedade. Sua atuação em decisões que afetam diretamente indivíduos e grupos sociais exige uma revisão crítica das normas éticas em vigor, destacando a importância de uma abordagem reflexiva e normativa diante de seu crescimento exponencial.

No entanto, Leal (2021) aponta que parte dos conflitos éticos mais relevantes decorre de um equívoco conceitual persistente: a confusão entre os termos “autônomo” e “automático”. Muitos dispositivos baseados em IA operam de forma independente, mas sem consciência ou responsabilidade moral, características exclusivas de agentes éticos humanos. Assim, rotular tais sistemas como “autônomos” é inadequado sob a ótica da filosofia moral. Tal distinção é essencial para o delineamento da responsabilidade ética e jurídica que envolve a criação e uso dessas tecnologias, fundamentando tanto a elaboração de normativas quanto a construção de posturas críticas e informadas por parte dos usuários e desenvolvedores.

Com o avanço tecnológico, novas questões éticas surgem, especialmente no campo da inteligência artificial (IA). A ética algorítmica preocupa-se com a aplicação responsável de sistemas automatizados, exigindo princípios como transparência, privacidade, responsabilidade e não discriminação. O uso ético da IA pressupõe que seus sistemas sejam compreensíveis e auditáveis, que protejam os dados pessoais e que não reforcem desigualdades existentes. Um dos principais riscos éticos está nos vieses algorítmicos, muitas vezes derivados de conjuntos de dados enviesados. Tais vieses podem reproduzir discriminações em áreas sensíveis como mercado de trabalho e acesso ao crédito. Para mitigar esse problema, Ribeiro, Eloi e Rodrigues (2025) afirmam que é fundamental implementar auditorias independentes, criar políticas de governança e assegurar a prestação de contas por possíveis falhas dos sistemas.

Nesse mesmo sentido, Kalluri (2021) propõe uma abordagem relacional à ética algorítmica, argumentando que o impacto social das tecnologias não pode ser avaliado apenas por métricas técnicas. Para o autor, é imprescindível considerar os contextos históricos, sociais e culturais, bem como ouvir as comunidades afetadas pelas decisões automatizadas. A justiça algorítmica, portanto, só será alcançada com uma participação mais ativa e inclusiva nos processos de desenvolvimento e regulação tecnológica.

Ribeiro, Eloi e Rodrigues (2025) destacam que, dentro dos princípios éticos aplicáveis à IA, a transparência ocupa papel central. Ela envolve a possibilidade de compreender e auditar os critérios utilizados pelos algoritmos, mesmo por pessoas sem formação técnica. Sua ausência compromete a legitimidade e dificulta a responsabilização por decisões automatizadas. Por isso, mecanismos de explicabilidade, rastreabilidade e controle social são fundamentais para garantir uma IA ética e confiável.

Outro princípio essencial é o da não discriminação. Gallent-Torres, Zapata-González e Ortego-Hernando (2023) enfatizam que esse princípio exige que os sistemas algorítmicos sejam desenvolvidos com dados representativos e livres de vieses que possam perpetuar injustiças estruturais, demandando uma preocupação ativa com equidade desde a concepção dos modelos.

Por fim, a *accountability* representa o compromisso ético e jurídico de que os responsáveis por projetar, implementar ou utilizar sistemas algorítmicos respondam por suas consequências. Isso significa que não pode haver zonas de impunidade diante de falhas automatizadas. Além do cumprimento legal, a *accountability* demanda uma postura ética proativa, voltada à prevenção de danos e à promoção do bem público, conforme Gonçalves e Barbaresco (2024).

3.3 O AR CABOUÇO REGULATÓRIO EUROPEU PARA A INTELIGÊNCIA ARTIFICIAL: RESTRIÇÕES, RISCOS E A IMPERATIVIDADE ÉTICA.

Enquanto o Brasil ainda debate os contornos de uma regulamentação específica para a Inteligência Artificial, a União Europeia, por meio do Regulamento (UE) 2024/1689, conhecido como "AI Act", estabeleceu um marco legal formal e abrangente que visa promover uma IA centrada no ser humano e de confiança, garantindo um elevado nível de proteção da saúde, segurança e direitos fundamentais. Este regulamento representa um passo significativo na governança da IA, adotando uma abordagem baseada no risco que proíbe certas práticas e impõe requisitos rigorosos a sistemas de alto risco, refletindo uma profunda preocupação ética em sua essência.

A legislação europeia proíbe expressamente uma série de práticas de IA consideradas inaceitáveis por violarem valores e direitos fundamentais da União. Entre elas, destacam-se a utilização de técnicas subliminares ou manipuladoras que possam distorcer o comportamento humano ou explorar vulnerabilidades, bem como sistemas de classificação social (social scoring) que levem a tratamentos prejudiciais em contextos não relacionados. Também são

vedadas avaliações de risco criminal baseadas exclusivamente em perfis, a criação aleatória de bases de dados de reconhecimento facial a partir da internet ou CCTV , e a inferência de emoções em ambientes de trabalho ou educação, exceto para fins médicos ou de segurança. Particularmente sensíveis são as proibições de sistemas de categorização biométrica que deduzem atributos sensíveis como raça ou orientação sexual, e o uso de identificação biométrica remota em tempo real em espaços públicos para fins de aplicação da lei, que é restrito a exceções estritamente necessárias e sujeitas a rigorosas salvaguardas e autorização prévia.

Além das proibições, o Regulamento categoriza sistemas de IA como de Alto Risco (Artigos 6º e 7º, e Anexo III), com base em seu potencial de causar danos significativos à saúde, segurança ou direitos fundamentais. Esses sistemas, que abrangem áreas como infraestruturas críticas, educação, emprego, serviços essenciais, aplicação da lei, gestão de migração e administração da justiça, estão sujeitos a requisitos obrigatórios rigorosos. Tais requisitos incluem a governança de dados para mitigar vieses e discriminação , documentação técnica abrangente , e a garantia de supervisão humana efetiva para que os operadores possam intervir e anular resultados problemáticos.

A dimensão ética perpassa todo o arcabouço regulatório europeu, visando que a IA seja "centrada no ser humano e de confiança". O regulamento incorpora os princípios das "Orientações Éticas para uma IA de Confiança" (AI HLEG), promovendo iniciativa e supervisão humanas, solidez técnica, privacidade, governança de dados, transparência, diversidade, não discriminação, equidade, bem-estar social e ambiental, e responsabilização. A transparência, em particular, é um pilar central, exigindo que os sistemas de alto risco sejam compreensíveis para seus utilizadores. Adicionalmente, o Regulamento prevê o direito à explicação para pessoas afetadas por decisões significativas tomadas com base em sistemas de IA de alto risco, reforçando a accountability.

Em contraste com esse panorama regulatório formal e detalhado na União Europeia, o Brasil ainda busca consolidar seu próprio arcabouço legal para a Inteligência Artificial. A ausência de uma regulamentação tão específica e abrangente quanto o AI Act da UE pode resultar em lacunas significativas na proteção de direitos, na gestão de riscos e na promoção de uma IA verdadeiramente ética e responsável no contexto da administração pública brasileira.

4 CONSIDERAÇÕES FINAIS

A análise deste artigo sobre os Modelos de Linguagem da Inteligência Artificial e seu uso na Administração Pública revela um cenário de oportunidades transformadoras, mas igualmente carregado de desafios éticos e jurídicos que exigem uma abordagem proativa e interdisciplinar. Iniciativas brasileiras como o "Poupinha" do INSS e o "Alice" do TCU ilustram o potencial desses modelos para aprimorar a eficiência administrativa e o atendimento ao cidadão, promovendo processos mais ágeis e acessíveis. Contudo, as limitações técnicas evidenciadas, como vieses algorítmicos, a opacidade das decisões ("caixa preta") e os imperativos de conformidade com a Lei Geral de Proteção de Dados (LGPD), sublinham que a integração da IA não pode comprometer os princípios constitucionais do artigo 37 da Constituição Federal, notadamente legalidade, impessoalidade e eficiência.

A reflexão ética, enraizada nas contribuições de Aristóteles e Kant e adaptada à ética algorítmica, emerge como um pilar essencial para orientar essa integração. A distinção entre "automático" e "autônomo" reforça que a IA, por mais avançada, carece de consciência moral, colocando a responsabilidade ética e jurídica nas mãos de seus criadores e usuários. Essa percepção crítica expõe a fragilidade da regulação brasileira atual, que contrasta com o rigor do AI Act europeu, que classifica sistemas por risco, exige transparência, gestão de dados anti-viés, supervisão humana e o direito à explicação. No contexto nacional, a ausência de um marco regulatório comparável amplifica vulnerabilidades, como discriminações estruturais e falta de accountability, especialmente diante da rápida evolução tecnológica.

Recentemente, o Projeto de Lei nº 2.338/2023, aprovado no Senado em dezembro de 2024, avança na Câmara dos Deputados desde março de 2025, com uma Comissão Especial liderada pelo Deputado Aguinaldo Ribeiro (PP-PB) instaurada em abril. Debates intensos, incluindo audiências públicas em junho e setembro de 2025, abordaram impactos setoriais (saúde, educação, trabalho), inovação e direitos autorais, com aprovação de requerimentos como o REQ 149/2025 (16/09/2025) para incluir associações em discussões sobre IA em serviços públicos. Apesar do progresso, o parecer do relator ainda está pendente, com ações como audiências adicionais e pensamentos de outros PLs (ex.: PL 1465/2024) em curso. Embora o PL enfatize uma IA centrada no humano, carece de disposições obrigatórias para auditorias éticas e sanções claras, evidenciando a necessidade de emendas que fortaleçam sua eficácia.

Diante disso, a implementação ética da IA na Administração Pública demanda uma governança robusta e colaborativa, reunindo juristas, cientistas de dados, sociedade civil e

policymakers para monitorar vieses em tempo real e adaptar a legislação à velocidade da inovação. A criação de comitês éticos permanentes nos órgãos públicos, inspirados no modelo do TCU, e a elaboração de um "AI Act Brasileiro" adaptado ao contexto nacional, com foco em privacy by design, classificação de riscos e participação pública, são passos urgentes. Além disso, a obrigatoriedade de relatórios auditáveis para decisões automatizadas, auditorias independentes de datasets com atenção à diversidade (gênero, raça, região), e a integração de sanções à LGPD para violações em IA pública podem mitigar os riscos identificados, alinhando o Brasil a padrões globais. Investimentos em capacitação de servidores e infraestrutura digital, especialmente em áreas rurais, inspirados no modelo da Estônia, também são indispensáveis para superar barreiras de inclusão e maximizar os benefícios da tecnologia.

O atraso na tramitação do PL 2.338/2023, ainda em debate na Comissão Especial, representa um risco significativo, dado o ritmo acelerado dos avanços em modelos generativos e a possibilidade de lacunas regulatórias serem exploradas. Acelerar sua aprovação, com emendas que reforcem auditorias obrigatórias e sanções proporcionais, é imperativo para evitar abusos e assegurar uma IA responsável. As implicações dessa integração vão além da esfera administrativa: uma IA bem regulamentada pode democratizar serviços públicos, reduzir burocracia e promover equidade, mas isso depende de princípios éticos sólidos e de uma governança adaptativa.

Por fim, este estudo sugere que o sucesso da IA na Administração Pública brasileira reside em sua capacidade de se alinhar aos valores democráticos e aos direitos fundamentais, transformando-a em uma ferramenta humanizada. Futuras pesquisas empíricas sobre os impactos da IA em contextos específicos, como saúde e educação, e diálogos internacionais para harmonizar regulamentações podem enriquecer essa trajetória. Assim, a tecnologia pode pavimentar um futuro onde eficiência, transparência e justiça convivam em equilíbrio, oferecendo um serviço público inclusivo e equitativo a todos os cidadãos brasileiros.

REFERÊNCIAS

ALTMAN, M. Explainable AI and Administrative Law: The Black Box Problem. *Harvard Journal of Law & Technology*, 2020.

ANGWIN, Julia. *et al.* Machine Bias: There's software used across the country to predict future criminals. And it's biased against blacks. **ProPublica**, 2016. Disponível em: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Acesso em: 10 abr. 2025.

BARBOSA, Lucas Romanowski. Perspectiva trágica na ética de Aristóteles. **PÓLEMONS - Revista de Estudantes de Filosofia da Universidade de Brasília**. v. 9, n. 18, p.12-30, 2020.

BENDER, Emily M. *et al.* On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? . In: PROCEEDINGS OF THE 2021 ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY, 2021. New York. Anais [...]. New York: Association for Computing Machinery, p. 610-623, 2021.

BIRD, Steven.; KLEIN, Ewan; LOPER, Edward. **Natural Language Processing with Python**: analyzing text with the natural language toolkit. Beijing: O'Reilly Media, 2009.
 BRASIL. [Constituição (1988)]. **Constituição da República Federativa do Brasil de 1988**. Brasília: Presidência da República, 1988. Disponível em:
http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 10 abr. 2025.

BRASIL. **Lei nº 9.784, de 29 de janeiro de 1999**. Regula o processo administrativo no âmbito da Administração Pública Federal. Brasília: DF, Presidência da República, 1999. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l9784.htm. Acesso em: 10 abr. 2025.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados Pessoais. Brasília: DF, Presidência da República, 2018. Disponível em:
http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 10 abr. 2025.

BRASIL. Senado Federal. **Projeto de Lei nº 2.338/2023**. Dispõe sobre o uso da Inteligência Artificial. Brasília: Senado Federal, 03 de maio de 2023. Disponível em:
<https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2338/2023>. Acesso em: 10 abr. 2025.

CALLEJÓN, F. Derecho Administrativo y Transformación Digital. Madrid: Civitas, 2020.

CANDIOTTO, Cesar. **Ética**: abordagens e perspectivas. 3. ed. Curitiba: PucPress, 2023.

DOWNEY, Allen B. **Think Python**: How to Think Like a Computer Scientist. 2. ed. Needham: Green Tea Press, 2015. Disponível em: <https://greenteapress.com/wp/think-python-2e/>. Acesso em: 10 abr. 2025.

FAJARDO, Ana María Fajardo. La ética de Immanuel Kant. **Revista Internacional De Filosofía Teórica Y Práctica**. Bogotá, v. 1. n. 1. Jan./Jun.2021. Disponível em:
<https://revistas.uexternado.edu.co>. Acesso em: 08 maio 2025.

FLORIDI, Luciano *et al.* AI4People-An Ethical Framework for a Good AI Society: Opportunities, Risks, and Recommendations. **Minds and Machines**, v. 30, n. 4, p. 689–707, 2020.

FLORIDI, L.; COWLS, J.; BELTRAMETTI, M.; et al. AI Governance and Ethics: The Role of Regulation. *AI & Society*, 2020.

GALLENT-TORRES, Cinta; ZAPATA-GONZÁLEZ, Alfredo; ORTEGO-HERNANDO, José Luis. The impact of Generative Artificial Intelligence in higher education: a focus on ethics and academic integrity. **RELIEVE - Revista Electrónica de Investigación y**

Evaluación Educativa,2023.

GARCIA, Ana Cristina Bicharra. Ética e Inteligência Artificial. In: Artificialmente Humanos ou Humanamente artificial?. **Revista da Sociedade Brasileira de Computação**. Rio de Janeiro, n. 43, 2020. p. 14-22. Disponível em: DOI: <https://doi.org/10.5753/combr.2020.43.1791>. Acesso em: 23 out. 2025.

GONÇALVES, Rafaela Vilela; BARBARESCO, Rogério Ananias. Discriminação em algoritmos de inteligência artificial: estudo da LGPD como mecanismo de controle dos vieses discriminatórios. **Revista Científica da UNIFENAS**, v. 6, n. 8, 2024.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep Learning**. MIT Press, 2016.

GOVTECH SINGAPORE. AI in Public Administration: Singapore Case Study, 2022. Disponível em: <https://www.govtech.gov.sg>. Acesso em: 10 abr. 2025.

TRANSFORMERS Documentation. **Hugging Face**. Disponível em: <https://huggingface.co/docs/transformers/>. Acesso em: 10 abr. 2025.

CASS, Stephen. The Top Programming Languages 2025. **IEEE Spectrum**, 2025. Disponível em: <https://spectrum.ieee.org/top-programming-languages>. Acesso em: 10 abr. 2025.

INGLADA-GALIANA, L.; CORRAL-GUDIÑO, L.; MIRAMONTES-GONZÁLEZ, P. Ethics and artificial intelligence. **Revista Clínica Española**. Barcelona, v. 224, n. 3, p. 178-186, 2024. DOI: <https://doi.org/10.1016/j.rceng.2024.02.003>. Acesso em: 23 out. 2025.

INSTITUTO NACIONAL DO SEGURO SOCIAL (INSS). Chatbot “Poupinha”. Disponível em: <https://www.gov.br/inss>. Acesso em: 10 abr. 2025.

JURAFSKY, Daniel; MARTIN, James H. **Speech and Language Processing**. 3rd ed. Stanford University: Draft, 2023.

KALLURI, Pratyusha. Algorithmic injustice: a relational ethics approach. **Patterns**, v. 2, n. 2, p. 100178, 2021. Disponível em: DOI: <https://doi.org/10.1016/j.patter.2021.100178>. Acesso em: 23 out. 2025.

MANNING, C. D. **Natural Language Processing with Python**. Stanford University, 2022. Disponível em: <https://web.stanford.edu/class/cs224n/>. Acesso em: 10 abr. 2025.

OGAR, Tom Eneji; BASSEY, Samuel Akpan. The implications of Kantian ethics for non-rational nature/beings: a study in environmental ethics. **International Journal of Environmental Pollution and Environmental Modelling - IJEPM**. v. 6, 2023. p. 93-103. RIBEIRO, Adriana; ELOI, Dárcio Luiz Costa; RODRIGUES, Renata Vicentini Monteiro. Transparência e ética no uso de IA generativa. **Revista Base Científica**, v. 3, 2025.

SANTOS, Antônio Carlos dos. Variações conceituais entre a ética e a moral. **Filosofia Unisinos**, São Leopoldo, v. 22, n. 2, 2021. p. 1-10.

TRIBUNAL DE CONTAS DA UNIÃO (TCU). Plataforma “Alice”. Disponível em:

<https://www.tcu.gov.br>. Acesso em: 10 abr. 2025.

UNIÃO EUROPEIA. Ethics Guidelines for Trustworthy AI. 2024. Disponível em: <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>. Acesso em: 10 abr. 2025.

UNIÃO EUROPEIA. General Data Protection Regulation (GDPR), 2016. Disponível em: <https://eur-lex.europa.eu/eli/reg/2016/679/oj>. Acesso em: 10 abr. 2025.

UNIÃO EUROPEIA. Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho de 13 de junho de 2024. Jornal Oficial da União Europeia, n. 1689, 2024. Disponível em: <http://data.europa.eu/eli/reg/2024/1689/oj>. Acesso em: 10 abr. 2025.

VASWANI, Ashish *et al.* Attention is All You Need. In: XXXI CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS, v. 30, 2017. Long Beach, **Anais** [...], Long Beach, 2017.

LEAL, Tatiana Dulima Zabala. La ética en inteligencia artificial desde la perspectiva del Derecho. **Via Inveniendi Et Iudicandi**, v. 16, n. 2, jul./dez. 2021. DOI: <https://doi.org/10.15332/19090528.6785>. Acesso em: 10 jun. 2025.

ZAHARIA, M.; CHEN, M.; CHOWDHURY, M.; et al. MLflow: A Platform for the Complete Machine Learning Lifecycle. Data Science Conference, 2018. Disponível em: <https://databricks.com/mlflow>. Acesso em: 10 jun. 2025.