



**SECRETARIA DE SEGURANÇA PÚBLICA
UNIVERSIDADE ESTADUAL DE GOIÁS – UEG
COORDENADORIA DE ENSINO
COORDENAÇÃO DE ENSINO PRESENCIAL E DE PÓS-GRADUAÇÃO
ESPECIALIZAÇÃO EM ALTOS ESTUDOS EM SEGURANÇA PÚBLICA**

PAULO AUGUSTO PEDROSO AMARAL

**INTELIGÊNCIA ARTIFICIAL COMO FERRAMENTA DE APOIO À REVISÃO
TÉCNICA DE LAUDOS: Uma análise dos riscos técnicos e jurídicos**

GOIÂNIA-GO

2025



PAULO AUGUSTO PEDROSO AMARAL

**INTELIGÊNCIA ARTIFICIAL COMO FERRAMENTA DE APOIO À REVISÃO
TÉCNICA DE LAUDOS: Uma análise dos riscos técnicos e jurídicos**

Artigo apresentado como exigência parcial para conclusão da disciplina Metodologia Científica do Curso Especialização em Altos Estudos em Segurança Pública (CAESP) pela Secretaria de Segurança Pública de Goiás e a Universidade do Estado de Goiás, sob a orientação do Prof. Me. Murilo Toscano de Carvalho.

GOIÂNIA-GO

2025

INTELIGÊNCIA ARTIFICIAL COMO FERRAMENTA DE APOIO À REVISÃO TÉCNICA DE LAUDOS: Uma análise dos riscos técnicos e jurídicos

ARTIFICIAL INTELLIGENCE AS A SUPPORT TOOL FOR TECHNICAL REVIEW OF FORENSIC REPORTS: An analysis of technical and legal risks

Aluno(a)¹: Paulo Augusto Pedroso Amaral
Orientador(a)²: Murilo Toscano de Carvalho

RESUMO

Resumo: A transformação digital na Administração Pública tem impulsionado a adoção da Inteligência Artificial (IA) como ferramenta de otimização, inclusive na Segurança Pública. No contexto da atividade pericial, a Inteligência Artificial surge como um poderoso assistente para a revisão de laudos, capaz de processar grandes volumes de dados e reduzir erros humanos. Contudo, a implementação dessa tecnologia, especialmente modelos opacos como as "caixas-pretas", apresenta riscos técnicos e jurídicos significativos. Este artigo tem como objetivo analisar os riscos de natureza técnica (opacidade e vieses) e jurídica (implicações da LGPD e responsabilidade do servidor) inerentes ao uso da Inteligência Artificial no apoio à revisão de laudos periciais, visando estabelecer parâmetros de governança para sua utilização segura. A pesquisa utiliza uma abordagem qualitativa, exploratória e descritiva, com procedimentos de revisão bibliográfica e documental. Os resultados indicam que, tecnicamente, a opacidade dos modelos é incompatível com a necessidade de explicabilidade da prova, e os vieses algorítmicos podem perpetuar discriminações, contaminando a análise. Juridicamente, destaca-se que o uso de ferramentas externas não homologadas para processar laudos viola o dever de sigilo imposto pelo Estatuto do Servidor de Goiás (Lei nº 20.756/2020), configurando infração grave. Conclui-se que a adoção da Inteligência Artificial é viável apenas mediante rigorosas diretrizes de governança, como a exigência de explicabilidade (*XAI*), auditorias de vies e a manutenção da supervisão humana qualificada (*human-in-the-loop*) como pilar central.

Palavras-chave: Inteligência Artificial; Laudos Periciais; Segurança Pública; Riscos Jurídicos; LGPD.

Abstract: The digital transformation in Public Administration has driven the adoption of Artificial Intelligence (AI) as an optimization tool, including in Public Security. In the context of forensic activities, Artificial Intelligence emerges as a powerful assistant for reviewing reports, capable of processing large volumes of data and reducing human error. However, the implementation of this technology, especially opaque models like "black boxes," presents significant technical and legal risks. This article aims to analyze the technical (opacity and bias) and legal (implications of the LGPD and public servant liability) risks inherent in using

¹ Perito Criminal de 1ª Classe, da Superintendência de Polícia Técnico Científica, engenheiro agrícola, especialista em Gerenciamento em Segurança Pública.

² Formado em ciências biológicas, Perito Criminal de 2ª classe, da Superintendência de Polícia Técnico Científica, mestre em Engenharia de produção pela Universidade Federal de Goiás.

Artificial Intelligence to support the review of forensic reports, intending to establish governance parameters for its safe use. The research employs a qualitative, exploratory, and descriptive approach, using bibliographic and documentary review procedures. The results indicate that, technically, the opacity of models is incompatible with the need for evidence explicability, and algorithmic biases can perpetuate discrimination, contaminating the analysis. Legally, it is highlighted that using non-homologated external tools to process reports violates the duty of confidentiality imposed by the Goiás State Public Servant Statute (Law No. 20.756/2020), constituting a serious infraction. It is concluded that the adoption of AI is viable only through rigorous governance guidelines, such as the requirement for explainability (XAI), bias audits, and the maintenance of qualified human supervision (human-in-the-loop) as a central pillar.

Keywords: Artificial Intelligence; Forensic Reports; Public Security; Legal Risks; Data Protection.

1. INTRODUÇÃO

A sociedade atual atravessa uma fase de acelerada transformação tecnológica, refletindo em todos os setores, manifestando-se proeminentemente na Administração Pública, onde hoje, a transformação digital não é apenas uma atualização tecnológica, e sim, uma necessidade para o Estado (Coronel; Sotelo, 2022). Segundo Aloísio (2021), no âmbito da Segurança Pública, a implementação de inovações tecnológicas é um passo fundamental para a otimização, gerando resultados rápidos e consistentes. Entre as principais tecnologias, é possível destacar a Inteligência Artificial, não como uma substituta do ativo humano, e sim como um assistente poderoso para o processamento de um maior volume de dados, bem como para a redução de erro humano (El-Din, 2022; Piraianu *et al.*, 2023).

No âmbito da segurança pública, o laudo pericial representa um pilar fundamental na persecução penal, materializando vestígios e fornecendo subsídio ao processo judicial. A confiabilidade, a qualidade, a segurança e a isenção do laudo pericial, envolve um processo rigoroso de revisão, que demandam tempo e recursos humanos significativos, sendo então esperado que um laudo influenciado por Inteligência Artificial mantenha a transparência e explicabilidade, bem como a validação e confiabilidade, sem descartar a responsabilidade e auditabilidade dos peritos criminais (Stacey *et al.*, 2025).

Nesse cenário, a Inteligência Artificial se destaca apresentando desafios que por ora, são pouco conhecidos, justamente por ser uma tecnologia nova. Por essa razão, o assunto demanda estudo para entender as reais dificuldades. O problema da “caixa-preta”, por exemplo, pode ser dividido em quatro problemas: a opacidade, que é a incapacidade de compreender o funcionamento interno e os mecanismos do sistema de Inteligência Artificial; o problema da estranheza, que é a dificuldade em se aceitar que um sistema pode tomar decisões complexas e autônomas; a imprevisibilidade das decisões; e a justificação, que é a dificuldade do sistema em justificar suas decisões (Brožek *et al.*, 2023).

A incapacidade de compreensão dos mecanismos internos inerente a muitos modelos de Inteligência Artificial, conhecidos como "caixas-pretas", pode ser um entrave para a previsibilidade e explicabilidade de suas decisões, podendo assim comprometer a sua adoção no processo penal, porém esses desafios podem ser gerenciáveis ao identificar os desafios reais do processo (Brožek *et al.*, 2023). Portanto, apesar de promissora tecnologia, o seu uso e adoção depende de análise dos desafios e das limitações, necessitando de esforço colaborativo entre seus principais entes (Anand; Thakur, 2025).

Com o avanço do uso da Inteligência Artificial no meio jurídico, problemas como o Duplo Potencial da Inteligência Artificial, que por um lado tem uma enorme capacidade de processamento de grande volume de dados, por outro, se não for devidamente compreendida e controlada, pode levar a decisões equivocadas, amplificando mais ainda as desigualdades sociais, necessitando assim de um debate crítico, respostas consolidadas e validação científica, em assuntos tais como a cadeia de custódia da prova, a responsabilidade do servidor e a validade de uma prova revisada por um algoritmo não transparente (Almada; Zanatta, 2024).

Este é o tema geral no qual a presente pesquisa se insere: a intersecção entre a Inteligência Artificial, a gestão da prova pericial e a Segurança Pública.

O uso da Inteligência Artificial pode otimizar os recursos empenhados e auxiliar na prevenção da violência, com o uso eficaz de dados, informações e conhecimento (Lu *et al.*, 2025; Telles, 2021). Contudo, a adoção de uma tecnologia tão poderosa e autônoma não pode ocorrer de forma acrítica (Vatamanu; Tofan, 2025). A passagem de um modelo de revisão puramente humano para um modelo híbrido, auxiliado por Inteligência Artificial, representa um desafio de gestão pública que vai muito além da mera implementação de um software. A dificuldade prática e de gestão está em como implementar uma ferramenta de Inteligência Artificial, muitas vezes operando como uma "caixa-preta", sem introduzir erros ou vieses que comprometam a integridade do laudo (Almada; Zanatta, 2024; Dwivedi *et al.*, 2019). O problema, portanto, não é meramente técnico, mas de gestão, pois cabe ao gestor público decidir sobre a adoção de uma tecnologia cujas implicações podem extrapolar sua capacidade de controle (Vatamanu; Tofan, 2025).

Diante desta dificuldade, a presente pesquisa é norteadada pela seguinte questão: Quais são os principais riscos de natureza técnica e jurídica associados à utilização de sistemas de Inteligência Artificial para a revisão de laudos periciais e que parâmetros de governança e gestão devem ser estabelecidos para assegurar a validade da prova e a conformidade legal e administrativa no âmbito da Segurança Pública?

O objetivo geral deste trabalho consiste em analisar os riscos de natureza técnica e jurídica inerentes à utilização da Inteligência Artificial como ferramenta de apoio à revisão de laudos periciais, de forma a estabelecer os parâmetros para sua utilização segura, responsável e em conformidade com o ordenamento jurídico vigente. Para alcançar este fim, foram traçados três objetivos específicos. Primeiramente, busca-se identificar o estado da arte das ferramentas de Inteligência Artificial aplicáveis à revisão de laudos e categorizar os riscos técnicos inerentes, como vieses e opacidade (Dwivedi *et al.*, 2019; Tynan, 2024). Em segundo lugar, analisa-se o arcabouço jurídico brasileiro aplicável, investigando as implicações da Lei Geral

de Proteção de Dados e as potenciais responsabilidades do servidor (Almada; Maranhão, 2023). Por fim, propõe-se diretrizes e parâmetros de governança para a implementação segura de tais ferramentas (Stacey *et al.*, 2025).

2 REVISÃO DA LITERATURA

2.1. Inteligência Artificial e Transformação digital na segurança pública

No contexto da Segurança Pública, a implementação de inovações tecnológicas é um passo fundamental para a otimização dos recursos, gerando resultados rápidos e consistentes (Aloísio, 2021). A Inteligência Artificial não é uma substituta do ativo humano, mas sim, um assistente poderoso para o processamento eficaz de um maior volume de dados e consequentemente contribui para a redução do erro humano (Chango *et al.*, 2024; Telles, 2021).

Segundo Paula (2025), a Inteligência Artificial oferece um potencial transformador nas operações policiais, porém deve ser implementada com responsabilidade ética, através de regulamentações específicas para sistemas de Inteligência Artificial na segurança pública, adoção de mecanismos confiáveis de transparência, incluindo auditoria independentes, e um modelo de governança participativo e transparente.

A Inteligência Artificial pode otimizar os recursos empenhados e auxiliar na prevenção da violência, com o uso eficaz de dados, informações e conhecimento (Lu *et al.*, 2025; Telles, 2021). A novidade da Inteligência Artificial é otimizar a análise de forma exponencial, numa escala com potencial de eficiência sem precedentes na história, dentro do contexto da Revolução Digital (Telles, 2021).

2.2. Aplicações da Inteligência Artificial nas Ciências Forenses

A utilização da Inteligência Artificial nas ciências forenses representa uma das mais significativas transformações na gestão da prova técnica, apresentando várias oportunidades de desenvolvimento tecnológico (Chango *et al.*, 2024). Ainda segundo Chango *et al.* (2024), a capacidade de processamento de dados em larga escala, a identificação de padrões complexos e a automação de tarefas repetitivas são promessas que podem redefinir a eficiência da atividade pericial. No entanto, o uso adequado da tecnologia de Inteligência Artificial requer o

reconhecimento de seus obstáculos técnicos e a devida compreensão de suas capacidades, o que é crucial para evitar a opacidade dos sistemas (Almada; Zanatta, 2024).

A Inteligência Artificial não é uma única tecnologia, ela abrange um espectro de tecnologias que vão desde sistemas simples baseados em regras até complexas redes neurais profundas (Stacey *et al.*, 2025). Na análise forense, sua aplicação tem se destacado em várias situações (El-Din, 2022). Autores como Anand e Thakur (2025) e Chango *et al.* (2024) demonstram o potencial da Inteligência Artificial em seus trabalhos, como o aumento da precisão dos peritos e a redução do tempo das análises forenses, possibilitando que estes consigam lidar com maiores volumes de dados e a realização de análises ainda mais complexas.

Ferramentas de Processamento de Linguagem Natural, por exemplo, podem ser treinadas para realizar a triagem e revisão de laudos textuais, bem como para a análise de sentimento, interpretando o tom de mensagens em análise de evidências baseadas em texto, a identificação de padrões em grandes conjuntos de dados textuais, para detecção de anomalias ou exceções, a automação de ferramentas forenses digitais, bem como geração de relatórios baseados em linguagem natural (Wickramasekara *et al.*, 2025; Yin *et al.*, 2025). Tais sistemas são capazes de verificar a consistência terminológica, assegurar que todos os quesitos formulados pela autoridade requisitante foram respondidos no laudo e comparar os achados com bancos de dados de casos anteriores, identificando anomalias ou padrões relevantes (Ketsekioulafis *et al.*, 2024). Tal como demonstrou Mandayam (2024), a Inteligência Artificial poderá ser empregada ainda para extração eficiente de dados e inclusive coleta automatizada em evidências digitais, análise de áudio de vídeo em casos que envolvem evidências multimídia, análise de tráfego de rede, detectando intrusões e triagem forense, para otimização e alocação de recursos humanos, acelerando as investigações.

O potencial de otimização é evidente. A automação da revisão preliminar permite que o perito criminal, cuja expertise é e continuará sendo insubstituível, foque seu tempo e esforço na análise crítica e interpretativa dos casos mais complexos, em vez de gastá-lo em verificações mecânicas (Fakiha, 2023). A IA, portanto, surge como uma ferramenta de apoio, que amplia a capacidade humana, porém sem poder substituí-la.

2.3. Desafios Técnicos da Inteligência Artificial na Atividade Pericial: “Caixa-Preta”

Apesar das potencialidades, a eficácia dessas ferramentas está intrinsecamente ligada à transparência de seus modelos. Segundo Anand e Thakur (2025), o primeiro grande risco

técnico da IA, especialmente dos modelos baseados em *deep learning*, é a sua opacidade, pois são complexos e difíceis para os humanos entenderem. Diversos autores, tais como Dwivedi *et al.* (2019), Kelly *et al.* (2020) e Durán *et al.* (2024), enfatizam que muitos dos sistemas mais avançados operam como "caixas-pretas" (*black boxes*), um termo utilizado para descrever algoritmos complexos nos quais o processo decisório interno não é plenamente compreensível nem mesmo para seus desenvolvedores, tornando a sua utilização em contextos sensíveis como a ciência forense, um desafio, já que falta de transparência, impede a compreensão, validação e confiança nas decisões da IA, levantando assim, preocupações éticas e legais.

Essa opacidade é fundamentalmente oposta à natureza da prova pericial, já que, entre as características de um laudo pericial, estão a clareza, objetividade e explicabilidade ou metodologia. O perito criminal deve fundamentar sua conclusão com base nos vestígios coletados. Se a revisão de um laudo é apoiada por uma Inteligência Artificial, e essa Inteligência Artificial sinaliza uma inconsistência crítica, mas o perito não consegue explicar por que o algoritmo chegou a esse resultado, a prova se torna vulnerável (Anand; Thakur, 2025; Durán *et al.*, 2024). O contraditório e a ampla defesa ficam irremediavelmente prejudicados se a defesa não pode questionar a metodologia empregada, simplesmente porque ela está oculta dentro de um código impenetrável (Sallavaci, 2025).

Diante disso, autores como Kelly *et al.* (2020) e Solanke (2022) têm buscado desenvolver o campo da "IA Explicável", ou XAI, no qual, o objetivo é criar modelos que sejam, por natureza, mais transparentes ou que possam fornecer justificativas claras para suas decisões. Para a atividade pericial, a adoção de sistemas "explicáveis" não é uma opção, mas uma condição fundamental para garantir a validade jurídica da prova.

2.4. Vieses Algorítmicos e Implicações Éticas na Produção da Prova

O segundo grande risco técnico é o viés algorítmico, ou também conhecido como *Bias* (Anand; Thakur, 2025). Autores como Joseph (2024), Xavier (2024) e Javed e Li (2025) enfatizam que um sistema de Inteligência Artificial não nasce neutro, ele é um reflexo direto dos dados com os quais foi alimentado e da equipe que o programou, assim, se uma Inteligência Artificial é treinada para revisar laudos utilizando uma base de dados histórica que contém vieses humanos, sejam eles sociais, raciais ou de gênero, o algoritmo não apenas aprenderá, mas poderá perpetuar e amplificar essas distorções em uma escala sem precedentes.

No contexto da segurança pública, o risco de racismo algorítmico é um problema já documentado, em que sistemas de policiamento preditivo ou reconhecimento facial demonstram taxas de erro desproporcionalmente maiores para grupos minoritários (Lopez, 2021; Vistisen; Jensen, 2020).

Transportando esse risco para a revisão de laudos periciais, a falta de diversidade nos dados de treinamento, como adverte Hofmann *et al.* (2024), pode contribuir para o racismo velado codificado nos dados de treinamento, manifestando-se nos modelos de linguagem. Assim como reforçado por Moore *et al.* (2023), um algoritmo pode ser preciso e justo em um local e não ser em outro. Tal cenário não é apenas uma falha técnica, é uma violação direta dos princípios de isonomia e justiça, contaminando a prova pericial em sua origem (Valdivia *et al.*, 2022, 2024). A proposta de Fährndrich *et al.* (2022) diante desses problemas é a seleção de dados com qualidade para o treinamento desses sistemas, bem como, fazer a coleta de dados de forma a reduzir o viés, e validando posteriormente com grande número de testes.

3. METODOLOGIA

A pesquisa caracteriza-se como aplicada, uma vez que busca oferecer subsídios concretos para o aprimoramento da gestão da prova pericial no âmbito da Superintendência de Polícia Técnico-Científica do Estado de Goiás. A abordagem adotada é qualitativa, voltada à análise e interpretação de conceitos, riscos e implicações jurídicas relacionadas ao uso de Inteligência Artificial (IA) na revisão técnica de laudos.

Em relação aos objetivos, o estudo se enquadra como exploratório e descritivo. É exploratório por examinar um tema recente e ainda pouco consolidado na prática pericial, permitindo maior familiaridade com os desafios tecnológicos e jurídicos envolvidos. É descritivo por identificar e caracterizar os riscos técnicos, como opacidade de modelos, ausência de explicabilidade e vieses algorítmicos, e os riscos jurídicos relacionados à proteção de dados e responsabilidade do servidor público.

Os procedimentos técnicos utilizados foram a pesquisa bibliográfica e a pesquisa documental. A pesquisa bibliográfica abrange artigos científicos com publicações nos últimos cinco anos anteriores, sobre Inteligência Artificial, governança algorítmica, ciências forenses e proteção de dados. A pesquisa documental contempla legislações aplicáveis, especialmente a Lei Geral de Proteção de Dados (LGPD), normas do legislativo pátrio, Estatuto do Servidor Público do Estado de Goiás (Lei nº 20.756/2020) e estudos relacionados ao uso de tecnologias emergentes no setor público.

A análise do material foi realizada por meio de raciocínio dedutivo, partindo-se de conceitos gerais de Inteligência Artificial, governança e vieses algorítmicos para, posteriormente, aplicar esses referenciais à realidade específica da revisão técnica de laudos.

4.1 RESULTADOS E DISCUSSÕES

Os resultados e discussões supervenientes se estruturam em três eixos temáticos, correlacionados diretamente com a atividade pericial: riscos técnicos, riscos jurídicos e diretrizes de governança. Ainda que não haja coleta de dados empíricos, os resultados são decorrência de análise conceitual aprofundada, com fundamentação especializada e de legislações positivadas do sistema legislativo nacional e do Estado de Goiás.

4.1. Riscos Técnicos Identificados no Uso de Inteligência Artificial na revisão de Laudos

A revisão da literatura evidencia que os dois principais riscos técnicos associados à utilização de sistemas de Inteligência Artificial na perícia criminal concentram-se na opacidade dos modelos de *deep learning* e nos vieses algorítmicos incorporados aos dados de treinamento. A opacidade diz respeito ao fato de que muitos desses modelos funcionam como verdadeiras “caixas-pretas”, nas quais o encadeamento lógico das decisões não é plenamente acessível nem mesmo aos seus desenvolvedores. Dwivedi *et al.* (2019), Kelly *et al.* (2020) e Durán *et al.* (2024) destacam que, em sistemas complexos, a relação entre entrada e saída é observável, mas o caminho interno percorrido pelo algoritmo permanece essencialmente inacessível, o que compromete a reprodutibilidade e a possibilidade de auditoria técnica das decisões. Em contexto pericial, essa característica é frontalmente incompatível com a exigência de transparência metodológica do laudo, na medida em que o perito criminal deve, quando requisitado, ser capaz de justificar suas discussões periciais no laudo.

Paralelamente, os vieses algorítmicos representam um segundo eixo crítico de risco. Joseph (2024), Xavier (2024) e Hofmann *et al.* (2024) demonstram que a Inteligência Artificial não é neutra: ela reflete, amplia e, em muitos casos, cristaliza distorções já presentes nos dados históricos utilizados em sua fase de treinamento. Em ambientes de segurança pública, isso significa que desigualdades estruturais de natureza social, racial ou territorial podem ser codificadas e reproduzidas de forma automatizada, gerando decisões que aparentam ser técnicas, mas que, na prática, carregam discriminações veladas. Na revisão de laudos periciais, esse fenômeno pode se traduzir em priorização indevida de determinados tipos de

inconsistência, em atribuição assimétrica de relevância a elementos probatórios ou mesmo em padrões de leitura do texto pericial que reforçam estereótipos, sem que o perito ou o gestor tenham plena consciência desse processo.

Esse conjunto de fatores compromete a confiabilidade da revisão técnica apoiada por Inteligência Artificial e introduz uma camada adicional de insegurança na cadeia de produção da prova. A combinação entre modelos opacos e bases de dados enviesadas cria um cenário em que erros ou distorções podem ser dificilmente detectados e facilmente naturalizados, já que são apresentados sob a aparência de objetividade algorítmica. Do ponto de vista da atividade pericial, isso exige que a adoção de sistemas de Inteligência Artificial seja acompanhada de critérios rigorosos de validação, testes de sensibilidade, auditorias de viés e definição clara de limites de uso, sob pena de se fragilizar não apenas a qualidade técnica do laudo, mas também a sua legitimidade perante o sistema de justiça.

4.2 Riscos Jurídicos e Institucionais: LGPD, Responsabilidade e Prova

A implementação de uma ferramenta de Inteligência Artificial na revisão de laudos periciais transcende o debate técnico e adentra também na seara jurídica. A gestão dessa tecnologia implica uma análise rigorosa de sua conformidade com o ordenamento pátrio, especialmente no que tange à proteção de dados pessoais e à definição de responsabilidade por eventuais falhas, seja na esfera cível, penal ou disciplinar. Segundo Sharma (2024), a coleta e análise de grandes volumes de dados por sistemas de IA, levantam preocupações significativas como a privacidade, assim como também é necessário o estudo da natureza autônoma das decisões impulsionadas por Inteligência Artificial, quanto à autoria da responsabilidade pelo erro.

Laudos periciais são fontes de dados pessoais e, não raro, de dados pessoais sensíveis, como informações biométricas, genéticas, de saúde ou sobre a vida sexual. Segundo Almada e Maranhão (2023), a utilização de um software de Inteligência Artificial para ler, processar e analisar esses documentos configura uma operação de tratamento de dados em larga escala, sob a luz da Lei nº 13.709/2018, a Lei Geral de Proteção de Dados Pessoais.

É fundamental notar que a LGPD, em seu Art. 4º, inciso III, excepciona o tratamento de dados realizado para fins exclusivos de segurança pública, defesa nacional, segurança do Estado ou atividades de investigação e repressão de infrações penais. Nesse sentido, princípios como o da finalidade, da necessidade, da prevenção e, sobretudo, o da segurança, são plenamente

aplicáveis. Contudo, conforme adverte Brandão e Brito (2023), essa exceção não representa uma carta branca para o uso indiscriminado da tecnologia. O mesmo artigo determina que tal tratamento deve ser regido por legislação específica, que, por sua vez, deve prever medidas proporcionais e estritamente necessárias ao atendimento do interesse público e observar os princípios gerais de proteção e os direitos do titular previstos nesta Lei.

A proteção de dados é considerada um direito fundamental em diversos contextos jurídicos (Shehu; Shehu, 2023). Ainda segundo Almada e Maranhão (2023), o gestor público que adota uma ferramenta de Inteligência Artificial assume o dever de garantir que o sistema possua barreiras técnicas robustas contra vazamentos, acessos não autorizados ou uso indevido das informações sensíveis contidas nos laudos. Permitir que dados sensíveis de cidadãos sejam processados por algoritmos opacos, sem garantias de segurança, representa uma grave violação desse dever de custódia.

A atribuição de responsabilidade por danos causados por sistemas de IA, como no caso de uma revisão falha do laudo pericial, é um desafio complexo que não se encaixa facilmente nos paradigmas legais existentes (Diamantis, 2020). Diante disso, Havu *et al.* (2024), enfatiza a necessidade de regras de responsabilidade civil mais claras adaptadas à IA.

Segundo a Constituição Federal, a responsabilidade civil do Estado, prevista no Art. 37, § 6º da Constituição Federal, é objetiva. Ou seja, o Estado responde pelo dano causado por seus agentes, independentemente de dolo ou culpa, bastando a comprovação do nexo causal entre a conduta e o dano. Se o laudo está errado por falha da ferramenta adotada pelo Estado, a responsabilidade primária de reparar o dano à vítima é estatal.

Na esfera penal, a situação é mais complexa, pois esta exige a comprovação de dolo ou culpa. Um perito poderia ser acusado de falso testemunho ou falsa perícia se agir com dolo, ou seja, se conscientemente utilizar a Inteligência Artificial para fraudar um resultado. Os autores Fotheringham e Smith (2024) e Fraser e Suzor (2025), sugerem que mesmo com a IA, a responsabilidade final do profissional humano permanece, portanto, o perito que confia cegamente em um resultado algorítmico, sem realizar as verificações devidas, poderia estar agindo com negligência ou imperícia.

A inserção de tais laudos periciais em plataformas de Inteligências Artificiais externas e não homologadas pelo órgão competente pode configurar um tratamento de dados ilícito, violando diretamente os princípios de responsabilidade e *accountability* pública (Sharma, 2024).

No que diz respeito à responsabilidade administrativa, o Estatuto dos Servidores Públicos Civis do Estado de Goiás (Lei nº 20.756/2020), estabelece como deveres do servidor

o zelo e a dedicação às atribuições do cargo, no artigo 192, I, bem como a observância às normas legais e regulamentares, no artigo 192, II. O uso de uma ferramenta de Inteligência Artificial não homologada pela instituição, ou a validação de uma revisão algorítmica de forma acrítica e negligente, pode configurar uma violação direta desses deveres. Se um erro da Inteligência Artificial não for detectado pelo perito criminal por falta de zelo, e isso comprometer a validade da prova, o servidor estará sujeito a um Processo Administrativo Disciplinar (Goiás, 2020). Cabe, portanto, à gestão pública, nesse sentido, o fornecimento de ferramentas seguras, e ao servidor, o dever de utilizá-las com a perícia que seu cargo exige (Nicolás; Sampaio, 2024).

Em nível estadual, as implicações legais para o uso indevido do laudo pericial são extensas. O Estatuto dos Servidores Públicos Cíveis do Estado de Goiás (Lei nº 20.756/2020), por exemplo, estabelece como dever funcional "abster-se de revelar informação sobre a qual deva guardar sigilo" (Art. 192, VI) e proíbe expressamente "revelar ou utilizar informação protegida por sigilo, da qual tem ciência em razão do cargo ou função, salvo nos casos autorizados por Lei" (Art. 202, LIX) (Goiás, 2020). Portanto, a utilização de uma Inteligência Artificial externa para processar laudos oficiais não representa uma inovação processual, mas sim uma infração disciplinar grave, passível de demissão por configurar revelação de sigilo (Goiás, 2020).

4.3. Diretrizes de Governança para Implementação Segura De IA

A gestão pública não pode abdicar da inovação, e deve, portanto, implementá-la de forma responsável, nesse contexto, Stacey *et al.* (2025), propõem diretrizes de governança baseadas na mitigação dos riscos identificados, assegurando que a tecnologia sirva como um instrumento de fortalecimento da justiça, e não de fragilização da prova.

A primeira diretriz é que nenhum sistema de Inteligência Artificial deve ser implementado sem um processo rigoroso de validação, auditoria e homologação (Lorch *et al.*, 2022). Isso exige que a instituição estabeleça critérios técnicos claros para a adoção de tais softwares (Stacey *et al.*, 2025).

Ainda, conforme propõe Stacey *et al.* (2025), os sistemas devem ter princípios fundamentais, tais como explicabilidade, confiança, segurança, justiça e considerações éticas. O principal critério deve ser a explicabilidade, pois sistemas baseados em caixas-pretas são incompatíveis com a atividade pericial, pois o processo de decisão interna não é totalmente compreensível (Durán *et al.*, 2024; Solanke, 2022). A instituição deve então priorizar ou exigir

modelos de Inteligência Artificial Explicável (XAI), em que o sistema seja capaz de fornecer rastros de auditoria claros, demonstrando como e por que chegou a determinada conclusão ou sinalizou uma inconsistência (Shamoo, 2024; Solanke, 2022).

Além disso, segundo Joseph (2024), a auditoria de viés deve ser uma etapa obrigatória, onde um sistema, antes da implementação, deve ser testado contra bases de dados controladas para garantir que ele não apresente vieses discriminatórios. Essa auditoria não deve ser única, mas sim periódica, assegurando que o sistema não aprenda vieses ao longo do tempo. É fundamental abordar a necessidade de auditorias regulares e transparentes para a detecção de vieses, pois a imparcialidade é um pilar da justiça.

Para garantir a conformidade legal e a segurança do servidor, algumas recomendações de gestão são fundamentais. Segundo Fakiha (2023), a mais importante é a manutenção da supervisão humana qualificada, conhecida como *human-in-the-loop*.

O sistema de Inteligência Artificial deve ser classificado, em qualquer protocolo interno, como uma ferramenta de apoio, e jamais como um decisor final, onde a responsabilidade pela assinatura do laudo e pela validação da revisão é, e deve continuar sendo, integralmente do perito criminal (Farber, 2025).

Em segundo lugar, segundo Stacey *et al.* (2025), é necessária a criação de Protocolos Operacionais Padrão específicos para o uso da IA, nos quais esses POPs devem definir claramente: em quais tipos de laudos a ferramenta pode ser usada, quais os procedimentos de verificação humana obrigatórios após a análise da Inteligência Artificial e como os resultados algorítmicos devem ser registrados no sistema de cadeia de custódia.

Por fim, Chango *et al.* (2024) ressaltam que é imperativo o treinamento contínuo, os peritos devem ser treinados para compreender os princípios básicos de funcionamento da ferramenta, suas limitações e, principalmente, seus tipos de erros mais comuns, no qual o perito tenha habilidade para saber quando confiar e, mais importante, quando desconfiar da sugestão algorítmica. Essas diretrizes de governança estão compiladas no Quadro 1

Quadro 1 – Diretrizes de Governança para Adoção de Inteligência Artificial no âmbito da segurança pública

Categoria	Pré-requisito / Diretriz Essencial
Processos Iniciais (Antes de Implementar)	Validação, Auditoria e Homologação: Exigência de um processo rigoroso para qualquer sistema de IA (Lorch <i>et al.</i>, 2022). Critérios Técnicos: A instituição deve definir critérios claros para a adoção (Stacey <i>et al.</i>, 2025). Auditoria de Viés: Deve ser uma etapa obrigatória e periódica para testar vieses discriminatórios (Joseph, 2024).
Requisitos do Sistema (O que a IA deve ter)	Explicabilidade (XAI): Critério principal. Sistemas "caixa-preta" são incompatíveis com a perícia (Durán et al., 2024; Solanke, 2022). Rastros de Auditoria: O sistema deve ser capaz de demonstrar como e por que chegou a uma conclusão (Shamoo, 2024; Solanke, 2022). Princípios Fundamentais: O sistema deve ser baseado em confiança, segurança, justiça e ética (Stacey <i>et al.</i>, 2025).
Diretrizes de Gestão (Como usar)	Supervisão Humana (<i>Human-in-the-loop</i>): A IA é uma ferramenta de apoio, jamais um decisor final (Fakiha, 2023; Farber, 2025). Protocolos Operacionais Padrão (POPs): Exige a criação de POPs específicos para o uso da IA (Stacey <i>et al.</i>, 2025). Treinamento Contínuo: Imperativo para que os peritos compreendam as limitações e os erros comuns da ferramenta (Chango <i>et al.</i>, 2024).

Fonte: Elaborado pelo autor (2025) com base na literatura revisada.

5 CONSIDERAÇÕES FINAIS

Este trabalho se propôs a analisar os riscos técnicos e jurídicos associados ao uso da Inteligência Artificial como ferramenta de apoio à revisão de laudos periciais, visando propor diretrizes para sua gestão segura no âmbito da Segurança Pública.

A pesquisa confirmou que, embora a Inteligência Artificial ofereça um potencial significativo para otimizar a atividade pericial, aumentando a celeridade e a capacidade de análise de grandes volumes de dados, sua implementação sem estudos apresenta riscos graves. Os riscos técnicos centrais identificados são a opacidade dos modelos baseados em "caixa-preta", que ferem os princípios do contraditório e da ampla defesa, e os vieses algorítmicos, que podem perpetuar discriminações e comprometer a isenção da prova.

Do ponto de vista jurídico, constatou-se que o arcabouço legal brasileiro, embora não tenha sido desenhado para a era da Inteligência Artificial, oferece os caminhos para sua regulação. A LGPD, mesmo com suas exceções para a segurança pública, exige a observância de princípios que impõem um dever de segurança e transparência no tratamento de dados sensíveis. A responsabilidade por falhas, embora complexa devido à tecnologia, recairá sobre o Estado e, em via de regresso, sobre o agente público que agir com negligência, imperícia ou imprudência ao validar um resultado algorítmico falho.

Respondendo ao problema de pesquisa, conclui-se que os parâmetros de governança essenciais para mitigar esses riscos são: a exigência de explicabilidade, vetando sistemas opacos, a realização de auditorias de viés periódicas, a manutenção da supervisão humana qualificada (*human-in-the-loop*) como regra inegociável e o desenvolvimento de protocolos claros, bem como o treinamento contínuo.

A Inteligência Artificial não deve ser vista como uma substituta do perito, mas como um instrumento de qualificação. Sua implementação segura exige uma gestão pública focada na transparência e na responsabilidade. Sugere-se, para trabalhos futuros, a análise de casos concretos onde a prova técnica apoiada por Inteligência Artificial tenha sido contestada judicialmente, a fim de verificar a recepção dessas tecnologias pelos tribunais brasileiros.

REFERÊNCIAS

- ALMADA, M.; MARANHÃO, J. S. de A. Contribuições e Limites da Lei Geral de Proteção de Dados para a Regulação da Inteligência Artificial no Brasil. **Direito Público**, v. 20, n. 106, 2023. Disponível em: <https://doi.org/10.11117/rdp.v20i106.6957>. Acesso em: set. 2025.
- ALMADA, M.; ZANATTA, R. A. F. Inteligência artificial, direito e pesquisa jurídica. **Revista USP**, São Paulo, Brasil, n. 141, p. 51–64, 2024. DOI: 10.11606/issn.2316-9036.i141p51-64. Disponível em: <https://revistas.usp.br/revusp/article/view/225209>. Acesso em: ago. 2025.
- ALOÍSIO, V. L. N. Intelligent Technology to Enhance Policing and Public Accountability. **Journal of Artificial Intelligence and Systems**, v. 3, n. 1, p. 48, 2021. Disponível em: <https://doi.org/10.33969/ais.2021.31004>. Acesso em: abr. 2025.
- ANAND, K.; THAKUR, S. Challenges and Limitations of AI in Forensic Science: A Critical Review. **International Journal of Research Publication and Reviews**, 2025. Disponível em: <https://doi.org/10.55248/gengpi.6.0125.0672>. Acesso em: ago. 2025.
- BRANDÃO, R.; BRITO, A. L. Research Note - Regulating Artificial Intelligence in Brazil: the case of facial recognition systems at the state level. **Observatório De La Economía Latinoamericana**, v. 21, n. 1, p. 373, 2023. Disponível em: <https://doi.org/10.55905/oelv21n1-021>. Acesso em: ago. 2025.
- BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil**. Brasília, DF: Senado Federal, 1988.
- BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Dispõe sobre a proteção de dados pessoais e altera a Lei nº 12.965, de 23 de abril de 2014 (Marco Civil da Internet). **Diário Oficial da União**: seção 1, Brasília, DF, p. 59, 15 ago. 2018.
- BROŽEK, B.; FURMAN, M.; JAKUBIEC, M.; KUCHARZYK, B. The black box problem revisited: Real and imaginary challenges for automated legal decision making. **Artificial Intelligence and Law**, v. 32, n. 2, p. 427, 2023. Disponível em: <https://doi.org/10.1007/s10506-023-09356-9>. Acesso em: ago. 2025.
- CHANGO, X.; FLOR-UNDA, O.; GIL-JIMÉNEZ, P.; GÓMEZ-MORENO, H. Technology in Forensic Sciences: Innovation and Precision. **Technologies**, v. 12, n. 8, p. 120, 2024. Disponível em: <https://doi.org/10.3390/technologies12080120>. Acesso em: set. 2025.
- CORONEL, P. L. H.; SOTELO, C. G. M. Transformación digital en la administración pública: desafíos para una gobernanza activa en el Perú. **Comunicación Revista de Investigación en Comunicación y Desarrollo**, v. 13, n. 2, p. 93, 2022. Disponível em: <https://doi.org/10.33595/2226-1478.13.2.594>. Acesso em: ago. 2025.
- DIAMANTIS, M. Algorithms Acting Badly: A Solution from Corporate Law. **SSRN Electronic Journal**, 2020. Disponível em: <https://doi.org/10.2139/ssrn.3545436>. Acesso em: set. 2025.

DURÁN, J. M.; VLOED, D. van der; RUIFROK, A.; YPMA, R. J. F. From understanding to justifying: Computational reliabilism for AI-based forensic evidence evaluation. **Forensic Science International Synergy**, v. 9, p. 100554, 2024. Disponível em: <https://doi.org/10.1016/j.fsisyn.2024.100554>. Acesso em: out. 2025.

DWIVEDI, Y. K. *et al.* Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. **International Journal of Information Management**, v. 57, p. 101994, 2019. Disponível em: <https://doi.org/10.1016/j.ijinfomgt.2019.08.002>. Acesso em: ago. 2025.

EL-DIN, E. A. A. ARTIFICIAL INTELLIGENCE IN FORENSIC SCIENCE: INVASION OR REVOLUTION? **Egyptian Society of Clinical Toxicology Journal**, v. 10, n. 2, p. 20, 2022. Disponível em: <https://doi.org/10.21608/escjt.2022.158178.1012>. Acesso em: ago. 2025.

FÄHNDRICH, J.; HONEKAMP, W.; POVALEJ, R.; RITTELMEIER, H.; BERNER, S. Special Issue on Application of AI in Digital Forensics. **KI - Künstliche Intelligenz**, v. 36, n. 2, p. 121, 2022. Disponível em: <https://doi.org/10.1007/s13218-022-00777-3>. Acesso em: ago. 2025.

FAKIHA, B. Enhancing Cyber Forensics with AI and Machine Learning: A Study on Automated Threat Analysis and Classification. **International Journal of Safety and Security Engineering**, v. 13, n. 4, p. 701, 2023. Disponível em: <https://doi.org/10.18280/ijss.130412>. Acesso em: jan. 2025.

FARBER, S. AI as a decision support tool in forensic image analysis: A pilot study on integrating large language models into crime scene investigation workflows. **Journal of Forensic Sciences**, 2025. Disponível em: <https://doi.org/10.1111/1556-4029.70035>. Acesso em: out. 2025.

FOTHERINGHAM, K.; SMITH, H. Accidental injustice: Healthcare AI legal responsibility must be prospectively planned prior to its adoption. **Future Healthcare Journal**, v. 11, n. 3, p. 100181, 2024. Disponível em: <https://doi.org/10.1016/j.fhj.2024.100181>. Acesso em: set. 2025.

FRASER, H.; SUZOR, N. Locating fault for AI harms: a systems theory of foreseeability, reasonable care and causal responsibility in the AI value chain. **Law Innovation and Technology**, p. 1, 2025. Disponível em: <https://doi.org/10.1080/17579961.2025.2469345>. Acesso em: nov. 2025.

GIL, A. C. **Como Elaborar Projetos de Pesquisa**. Editora Atlas. 4 ed. São Paulo: Atlas, 2002.

GOIÁS (Estado). Lei nº 20.756, de 28 de janeiro de 2020. Dispõe sobre o regime jurídico dos servidores públicos civis do Estado de Goiás, das autarquias e fundações públicas estaduais, e dá outras providências. **Diário Oficial do Estado**, Goiânia, 29 jan. 2020.

HAVU, K. *et al.* Regulating Liability for AI-Induced Harm: Developments in EU Law and Insights from a Research Project. **Helsinki Legal Studies Research Paper No. 87**, 2024. Disponível em: <https://doi.org/10.2139/ssrn.4861450>. Acesso em: nov. 2025.

HOFMANN, V. *et al.* AI generates covertly racist decisions about people based on their dialect. **Nature**, v. 633, n. 8028, p. 147, 2024. Disponível em: <https://doi.org/10.1038/s41586-024-07856-5>. Acesso em: out. 2025.

JAVED, K.; LI, J. Bias in adjudication: Investigating the impact of artificial intelligence, media, financial and legal institutions in pursuit of social justice. **PLoS ONE**, v. 20, n. 1, 2025. Disponível em: <https://doi.org/10.1371/journal.pone.0315270>. Acesso em: set. 2025.

JOSEPH, J. Predicting crime or perpetuating bias? The AI dilemma. **AI & Society**, 2024. Disponível em: <https://doi.org/10.1007/s00146-024-02032-9>. Acesso em: jul. 2025.

KELLY, L. *et al.* **Explainable Artificial Intelligence for Digital Forensics: Opportunities, Challenges and a Drug Testing Case Study**. [S. l.]: IntechOpen, 2020. E-book. Disponível em: <https://doi.org/10.5772/intechopen.93310>. Acesso em: ago. 2025.

LOPEZ, P. Bias does not equal bias: a socio-technical typology of bias in data-based algorithmic systems. **Internet Policy Review**, v. 10, n. 4, 2021. Disponível em: <https://doi.org/10.14763/2021.4.1598>. Acesso em: jul. 2025.

LORCH, B.; SCHELER, N.; RIESS, C. Compliance Challenges in Forensic Image Analysis Under the Artificial Intelligence Act. In: **2021 29th European Signal Processing Conference (EUSIPCO)**. [S. l.: s. n.], 2022. p. 613. Disponível em: <https://doi.org/10.23919/eusipco55093.2022.9909723>. Acesso em: set. 2025.

LU, H. *et al.* Lightweight deep learning model for crime pattern recognition based on transformer with simulated annealing sparsity and CNN. **Scientific Reports**, v. 15, n. 1, 2025. Disponível em: <https://doi.org/10.1038/s41598-025-07260-7>. Acesso em: set. 2025.

MANDAYAM, R. The Impact of Artificial Intelligence on Digital Forensic. **Journal of Artificial Intelligence & Cloud Computing**, v. 3, n. 6, p. 1, 2024. Disponível em: [https://doi.org/10.47363/jaicc/2024\(3\)414](https://doi.org/10.47363/jaicc/2024(3)414). Acesso em: out. 2025.

MOORE, C.; FERGUSON, E.; GUERIN, P. Pretrial Risk Assessment on the Ground: Algorithms, Judgments, Meaning, and Policy. **MIT Case Studies in Social and Ethical Responsibilities of Computing**, Summer 2023. Disponível em: <https://doi.org/10.21428/2c646de5.b016a7b3>. Acesso em: out. 2025.

NICOLÁS, M. A.; SAMPAIO, R. C. Balancing efficiency and public interest: The impact of AI automation on social benefit provision in Brazil. **Internet Policy Review**, v. 13, n. 3, 2024. Disponível em: <https://doi.org/10.14763/2024.3.1799>. Acesso em: out. 2025.

PAULA, L. C. de. MART SAMPA: BETWEEN EFFICIENCY AND ETHICS -A CRITICAL ANALYSIS OF AI USE IN SÃO PAULO'S PUBLIC SECURITY. **Revista Científica da Polícia Militar de Mato Grosso do Sul**, v. 2, n. 2, p. 188, 2025. Disponível em: <https://doi.org/10.62927/revpmms.v2i2.103>. Acesso em: set. 2025.

PIRAIANU, A. I. *et al.* Enhancing the Evidence with Algorithms: How Artificial Intelligence Is Transforming Forensic Medicine. **Diagnostics**, 2023. Disponível em: <https://doi.org/10.3390/diagnostics13182992>. Acesso em: out. 2025.

SALLAVACI, O. Algorithms on Trial: Does evaluative probabilistic reporting of forensic evidence infringe the presumption of innocence? **Forensic Science International Synergy**, v. 11, p. 100591, 2025. Disponível em: <https://doi.org/10.1016/j.fsisyn.2025.100591>. Acesso em: ago. 2025.

SAMPAIO, R. C. *et al.* Muita Bardin, pouca qualidade: uma avaliação sobre as análises de conteúdo qualitativas no Brasil. **Revista Pesquisa Qualitativa**, v. 10, n. 25, p. 464, 2022. Disponível em: <https://doi.org/10.33361/rpq.2022.v.10.n.25.547>. Acesso em: jul. 2025.

SHAMOO, Y. The Role of Explainable AI (XAI) in Forensic Investigations. *In: Advances in digital crime, forensics, and cyber terrorism book series*. [S. l.]: IGI Global, 2024. E-book. p. 31. Disponível em: <https://doi.org/10.4018/979-8-3373-0857-9.ch002>. Acesso em: out. 2025.

SHARMA, N. ARTIFICIAL INTELLIGENCE: LEGAL IMPLICATIONS AND CHALLENGES. **Knowledgeable Research A Multidisciplinary Journal**, v. 2, n. 11, p. 13, 2024. Disponível em: <https://doi.org/10.57067/220k4298>. Acesso em: set. 2025.

SHEHU, V. P.; SHEHU, V. Human rights in the technology era - Protection of data rights. **European Journal of Economics, Law and Social Sciences**, v. 7, n. 2, p. 1, 2023. Disponível em: <https://doi.org/10.2478/ejels-2023-0001>. Acesso em: set. 2025.

SOLANKE, A. A. Explainable digital forensics AI: Towards mitigating distrust in AI-based digital forensics analysis using interpretable models. **Forensic Science International Digital Investigation**, v. 42, p. 301403, 2022. Disponível em: <https://doi.org/10.1016/j.fsidi.2022.301403>. Acesso em: out. 2025.

STACEY, J. *et al.* A Responsible Artificial Intelligence Framework for Forensic Science. **Forensic Science International**, v. 375, p. 112548, 2025. Disponível em: <https://doi.org/10.1016/j.forsciint.2025.112548>. Acesso em: ago. 2025.

TELLES, P. S. B. Inteligência artificial e polícia preditiva. **Boletim Científico Escola Superior do Ministério Público da União**, n. 57, p. 247, 2021. Disponível em: <https://doi.org/10.63601/bcesmpu.2021.n57.247-263>. Acesso em: maio 2025.

TYNAN, P. The integration and implications of artificial intelligence in forensic science. **Forensic Science Medicine and Pathology**, v. 20, n. 3, p. 1103, 2024. Disponível em: <https://doi.org/10.1007/s12024-023-00772-6>. Acesso em: set. 2025.

VALDIVIA, A.; HYDE-VAAMONDE, C.; GARCÍA-MARCOS, J. Judging the algorithm: A case study on the risk assessment tool for gender-based violence implemented in the Basque country. **arXiv (Cornell University)**, 2022. Disponível em: <https://doi.org/10.48550/arxiv.2203.03723>. Acesso em: out. 2025.

VALDIVIA, A.; HYDE-VAAMONDE, C.; MARCOS, J. Judging the algorithm. **AI & Society**, 2024. Disponível em: <https://doi.org/10.1007/s00146-024-02016-9>. Acesso em: out. 2025.

VATAMANU, A. F.; TOFAN, M. Integrating Artificial Intelligence into Public Administration: Challenges and Vulnerabilities. **Administrative Sciences**, v. 15, n. 4, p. 149, 2025. Disponível em: <https://doi.org/10.3390/admsci15040149>. Acesso em: ago. 2025.

VISTISEN, P.; JENSEN, T. The Selfish Vision: Addressing the problematic lack of explicit discourse and ethical contract in corporate design fictions. **Paradigm Shifts in ICT Ethics: Proceedings of the ETHICOMP 2020 18th International Conference on the Ethical and Social Impacts of ICT** (18 ed., Vol. 1, pp. 385-388). Universidad de La Rioja. Disponível em: <https://vbn.aau.dk/en/publications/the-selfish-vision-addressing-the-problematic-lack-of-explicit-di/>. Acesso em: jul. 2025.

XAVIER, B. Biases within AI: challenging the illusion of neutrality. **AI & Society**, 2024. Disponível em: <https://doi.org/10.1007/s00146-024-01985-1>. Acesso em: ago. 2025.