Coordenador
Higor Vinicius Nogueira Jorge

INTELIGÊNCIA ARTIFICIAL, DIREITO PENAL E JUSTIÇA CRIMINAL

Teoria, Garantias Fundamentais e Governança Algorítmica

Autores

- Amanda Maria Mota Tavares
- Beatriz Souto Orengo
- Bianca Branco
- Carlos Afonso Gonçalves da Silva
- Claudia da Costa Bonard de Carvalho
- Higor Vinicius Nogueira Jorge
- Líbero Carvalho
- Luiz Otávio Sales Damasceno
- Patrícia Soares Godoy
- Paula Moura da Silva Dias
- Roberta Cristiane Rocha
- Walter Martins Muller

2026

EDITORA JusPODIVM
www.editorajuspodivm.com.br

CAPÍTULO 12

GOVERNANÇA EM COMPLIANCE DIGITAL PARA INTELIGÊNCIA ARTIFICIAL

Claudia da Costa Bonard de Carvalho

INTRODUÇÃO

A inteligência artificial está redefinindo o cenário do mundo dos negócios e a realidade das empresas, onde muitas apenas aplicam, de forma parcimoniosa, tais tecnologias, seja pela demora no retorno sobre o investimento ou pela dificuldade de integrá-la aos seus processos, mas, ao mesmo tempo, outras querem mergulhar nesse mar de possibilidades novas que podem ser experimentadas.

Nesse último caso, temos empresas que querem remodelar seu *business* agora com agentes inteligentes, que já caracterizam a terceira onda da inteligência artificial, de forma a obter melhores resultados em diversos assuntos, como marketing, experiência do cliente, vendas e etc.

No caso, esse *rebranding* de atividades faz com que as empresas tenham que encarar outros desafios na implementação de qualquer forma de inteligência artificial, dadas as novas vulnerabilidades que

possam surgir em seus sistemas e crescentes ameaças contra bancos de dados.

Para tanto, há que se repensar a governança digital da empresa, estabelecendo-se agora também uma modalidade especial, relacionada à gestão da inteligência artificial, a qual requer diretrizes bem definidas, de modo a ser desempenhada da maneira mais eficiente, assegurando-se resultados satisfatórios em seus projetos de inovação.

Neste sentido, há diversos normativos que podem guiar inúmeros aspectos da governança de inteligência artificial, considerando-se as suas dificuldades e os compromissos com a conformidade digital que as empresas necessitam assumir ao utilizar estas tecnologias, devido aos seus riscos e impactos, os quais irão orientar gestores em suas atividades, e serão estudados a seguir.

12.1. OBJETIVOS DE GOVERNANÇA DE AI

A adoção de novas tecnologias exige que o board das organizações assuma um papel ativo e estratégico na avaliação das consequências do uso de soluções de inteligência artificial, de modo que, a responsabilidade de gestores não se limita à meros aspectos da implementação técnica, pois a sua negligência pode comprometer todo o sucesso de um projeto.

Logo, a gestão eficaz da inteligência artificial requer uma governança capaz de mitigar riscos e garantir que a tecnologia esteja alinhada aos objetivos estratégicos da empresa, para justificar a alocação de recursos em soluções a serem adquiridas no mercado.

A norma ISO/IEC 38507:2022 oferece diretrizes fundamentais para esse processo, destacando a importância de definir-se claramente os riscos associados à IA, bem como as obrigações adicionais relacionadas à sua manutenção e supervisão.

Essa norma orienta os membros do board, gestores e stakeholders sobre como integrar a inteligência artificial ao negócio de forma estratégica e segura, promovendo uma cultura organizacional que valorize a responsabilidade e a conformidade, conforme seu item 4.2:

> **4.2. Maintaining governance when introducing AI**
>
> The governing body sets the purpose of the organization and approves the strategies necessary to achieve that purpose.

> However, it is possible that existing governance is no longer fit-for-purpose when AI is being used within that organization. The specific choice of tools, e.g. AI systems, should be a management decision, made in light of and in line with guidance from the governing body. In order to establish such guidance, the governing body should inform itself about AI in general terms because its use can bring:
>
> — significant benefit to the organization strategically;
>
> — significant risk to the organization, with the potential for harm to its stakeholders;
>
> — additional obligations to the organization.

Nesse sentido, a escolha das ferramentas de inteligência artificial é uma decisão de gestão crítica e deve ser coerente com as metas que se pretende alcançar, pelo que, deve-se considerar não apenas o seu desempenho técnico, mas também aspectos como segurança da informação, privacidade de dados e capacidade de adaptação às necessidades do negócio.

Além disso, é essencial verificar a robustez dos controles de segurança da informação, protegendo-se os sistemas e dados contra ataques cibernéticos e acessos indevidos.

Não bastasse isso, o board deve assumir a responsabilidade direta pela governança de inteligência artificial, compreendendo os riscos envolvidos e promovendo uma supervisão contínua, o que inclui o estabelecimento de políticas claras, definição de métricas de desempenho e monitoramento constante dos resultados, através de auditorias.

O reporte transparente para stakeholders sobre o desempenho da inteligência artificial também é igualmente importante, permitindo que investidores, clientes e parceiros compreendam os benefícios e limitações da tecnologia adotada.

Outro aspecto essencial é a revisão periódica dos critérios de monitoramento da inteligência artificial devido à possibilidade de alucinações (respostas incorretas ou enganosas geradas por seus sistemas), pelo que, é necessário atualizar constantemente os parâmetros de avaliação, garantindo que a tecnologia opere dentro dos limites aceitáveis de precisão e confiabilidade.

Em suma, a governança da inteligência artificial deve ser tratada como um componente fundamental da gestão empresarial, ou seja,

o envolvimento ativo do board, a escolha criteriosa das ferramentas, a gestão dos riscos e a comunicação transparente com stakeholders são pilares fundamentais para o bom aproveitamento de ferramentas de inteligência artificial.

12.2. GOVERNANÇA E GESTÃO DE RISCO DE INTELIGÊNCIA ARTIFICIAL

Assim sendo, a implementação de projetos de inteligência artificial nas organizações deve ser precedida por uma análise rigorosa dos riscos já mapeados nas operações da empresa. Essa análise pode ser conduzida por um comitê especial do board, cuja função é avaliar se a introdução da tecnologia pode ampliar vulnerabilidades existentes, comprometendo a segurança, a conformidade ou a eficiência dos processos internos.

O objetivo da gestão especial de riscos de inteligência artificial é preparar a empresa para lidar com os novos riscos que podem ser introduzidos pela aplicação dessas ferramentas, sendo que, esses riscos não se limitam aos aspectos técnicos da implementação da tecnologia (sobrecarga de fluxo de dados, arquitetura adequada de sistemas de segurança da informação), mas abrangem também implicações éticas, jurídicas, operacionais e reputacionais de sua aplicação, exigindo uma abordagem integrada e estratégica.

Os parâmetros da gestão de risco de inteligência artificial devem então estar inseridos nas políticas de Compliance Digital da empresa, que devem ser definidas e aprovadas pelo board, de forma que, essas políticas devem estabelecer diretrizes claras para o uso responsável destas soluções, incluindo critérios de avaliação, monitoramento contínuo, e mecanismos de responsabilização.

Nesse contexto, a norma NIST AI Risk Management Framework (RMF 1) surge como referência para a definição das melhores práticas de gestão de risco aplicáveis aos sistemas de inteligência artificial. Essa norma fornece um conjunto estruturado de princípios e processos que podem ser incorporados às políticas corporativas, promovendo maior segurança e confiabilidade na adoção da tecnologia, conforme seu item 1.2.3:

1.2.3. Risk Prioritization

Attempting to eliminate negative risk entirely can be counterproductive in practice because not all incidents and failures can be eliminated. Unrealistic expectations about risk may lead organizations to allocate resources in a manner that makes risk triage inefficient or impractical or wastes scarce resources. A risk management culture can help organizations recognize that not all AI risks are the same, and resources can be allocated purposefully.

Actionable risk management efforts lay out clear guidelines for assessing trustworthiness of each AI system an organization develops or deploys. Policies and resources should be prioritized based on the assessed risk level and potential impact of an AI system. The extent to which an AI system may be customized or tailored to the specific context of use by the AI deployer can be a contributing factor..

O objetivo central da norma é evitar os impactos negativos da inteligência artificial nas rotinas dos seus usuários, por meio da adoção de uma cultura organizacional voltada à gestão de risco, de modo que, tal cultura deve ser disseminada em todos os níveis da empresa, garantindo-se que os riscos sejam identificados, avaliados, mitigados e monitorados de forma contínua.

Os pontos centrais de processos da norma NIST AI RMF são: mapeamento, medição, gerenciamento e governança de riscos, sendo que, estes elementos funcionam de forma interdependente, em um sistema semelhante ao ciclo PDCA (Plan-Do-Check-Act), onde as etapas iniciais são fundamentais para a eficácia desta governança.

O mapeamento permite identificar os riscos novos no ambiente de inteligência artificial, considerando o contexto específico de funcionamento dos sistemas, sendo que, essa etapa exige uma visão global de todas as partes envolvidas, incluindo desenvolvedores, usuários, fornecedores e reguladores, para garantir que os riscos sejam compreendidos em sua totalidade.

Sobre a medição dos riscos, esta envolve análises quantitativas e qualitativas das funcionalidades dos sistemas de inteligência artificial, pelo que, esta avaliação deve ser devidamente documentada e baseada em testes de performance, que revelem o comportamento das tecnologias em diferentes cenários operacionais e seus respectivos resultados.

O gerenciamento dos riscos também requer a alocação de recursos adequados para tratar os riscos já mapeados e medidos, tendo em vista que, essa alocação deve estar alinhada aos objetivos definidos pela governança corporativa, priorizando ações que reduzam a exposição da empresa a falhas, abusos ou impactos negativos.

Por fim, a governança exige a implementação de uma cultura organizacional em seus aspectos práticos, voltada à gestão de risco de inteligência artificial, alinhada aos princípios institucionais e às estratégias de negócios. Isso implica estruturar processos internos, manter documentação atualizada e antecipar cenários futuros de risco, promovendo uma abordagem proativa e resiliente frente aos desafios da inteligência artificial, conforme o item 5.1 da NIST AI RMF:

> **5.1 Govern**
>
> GOVERN is a cross-cutting function that is infused throughout AI risk management and enables the other functions of the process. Aspects of GOVERN, especially those related to compliance or evaluation, should be integrated into each of the other functions. Attention to governance is a continual and intrinsic requirement for effective AI risk management over an AI system's lifespan and the organization's hierarchy.
>
> Strong governance can drive and enhance internal practices and norms to facilitate organizational risk culture. Governing authorities can determine the overarching policies that direct an organization's mission, goals, values, culture, and risk tolerance.

12.3. TRANSPARÊNCIA E EXPLICABILIDADE PARA INTELIGÊNCIA ARTIFICIAL

Na medida que os sistemas inteligentes se tornam mais complexos e autônomos, cresce também a necessidade de garantir que os resultados gerados por essas tecnologias sejam supervisionados, compreendidos e validados por seres humanos, antes de sua aplicação em empresas, mitigando-se eventuais falhas nas suas conclusões

Assim sendo, o uso de inteligência artificial requer uma supervisão humana contínua, que assegure que os resultados estejam alinhados com os objetivos éticos, legais e estratégicos das instituições, o que faz parte das boas práticas de sua governança digital.

Logo, a supervisão humana tem por objetivo a análise crítica das suas conclusões, de modo a contextualizar a informação e identificar potenciais enviesamentos ou erros sistémicos, que possam ter consequências danosas.

Um dos pilares dessa supervisão é o entendimento sobre como a inteligência artificial alcançou determinada conclusão, de modo que, em sistemas baseados em *machine learning (*uso de algoritmos para organizar dados, reconhecer padrões e fazer com que computadores aprendam com esses modelos para gerar insights inteligentes sem a necessidade de pré-programação)[1] ou *deep learning (*é a parte do aprendizado de máquina que, por meio de algoritmos de alto nível, imita a rede neural do cérebro humano)[2] , o processo de decisão nem sempre é transparente, o que pode gerar dúvidas quanto à origem e à validade das respostas apresentadas.

Para contornar essa dificuldade, surge o conceito de "*Reasoning*", ou raciocínio da máquina — uma técnica que procura compreender o processo lógico subjacente ao funcionamento dos modelos linguísticos de larga escala (LLMs):

Reasoning in artificial intelligence (AI) refers to the mechanism of using available information to generate predictions, make inferences and draw conclusions. It involves representing data in a form that a machine can process and understand, then applying logic to arrive at a decision.[3]

Este tipo de análise permite aos especialistas examinar as inferências, as relações de causa e efeito e as sequências de pensamento que levaram a determinada resposta, tornando o sistema mais compreensível e confiável.

Desta forma, percebemos que a confiabilidade dos resultados da inteligência artificial tem um impacto direto nos processos decisórios empresariais e institucionais, tendo em vista que, um gestor que baseia as suas decisões em recomendações geradas por inteligência artificial precisa estar seguro de que essas informações são precisas, auditáveis e fundamentadas em dados de qualidade.

---

1. MACHINE LEARNING. Disponível em: https://www.salesforce.com/br/blog/machine-learning-vs-deep-learning/. Acesso em: 17 out. 2025.
2. DEEP LEARNING. Disponível em: https://www.salesforce.com/br/blog/machine-learning-vs-deep-learning/. Acesso em: 21 out. 2025.
3. REASONING. Disponível em: https://www.ibm.com/think/topics/ai-reasoning. Acesso em: 21 out. 2025.

Um erro não supervisionado pode provocar consequências graves, como perdas financeiras, ou danos à reputação da empresa e violações legais, pelo que, a confiança nos sistemas de inteligência artificial depende de uma combinação entre tecnologia robusta e supervisão humana eficaz.

Neste contexto, destaca-se a norma ISO/IEC 24028:2020, que estabelece diretrizes sobre transparência, confiabilidade e explicabilidade em sistemas de inteligência artificial, uma vez que, esta norma propõe diretrizes para assegurar que as decisões tomadas por algoritmos possam ser rastreadas, compreendidas e auditadas, conforme seu item 5.1:

> **5.1 Background**
>
> For the purposes of this document, it is important to provide working definitions of artificial intelligence (AI) systems and trustworthiness.
>
> Here, we consider an AI system to be any system (whether a product or a service) that uses AI. There are many different kinds of AI systems. Some are implemented completely in software, while others are mostly implemented in hardware (e.g. robots).
>
> A working definition of trustworthiness is the ability to meet stakeholders' expectations in a verifiable way. This definition can be applied across the broad range of AI systems, technologies and application domains..

Com isso, cria-se um ambiente mais seguro e previsível para a adoção da inteligência artificial em setores como finanças, saúde, educação e administração pública.

Ao promover a transparência, aquela norma facilita também a rastreabilidade de decisões, permitindo reconstruir o caminho seguido pelo sistema até à obtenção de determinado resultado, sendo que, essa rastreabilidade é essencial para fins de auditoria, compliance e governança digital.

Em matéria de Compliance Digital e transparência, um dos aspetos mais relevantes é a documentação dos dados e dos modelos utilizados, pois tal prática permite que cada versão de um modelo de inteligência artificial seja devidamente registrada, com informações sobre as suas fontes de dados, parâmetros, limitações e métricas de desempenho.

Essa documentação garante não só a rastreabilidade técnica, mas também o cumprimento de standards regulatórios, bem como reforça a responsabilidade das organizações quanto ao uso transparente das ferramentas de inteligência artificial.

12.4. ÉTICA E RESPONSABILIDADE

A ética no uso da inteligência artificial nas empresas também é um tema cada vez mais relevante, por conta dos novos riscos já mencionados, que podem surgir, diante da falta de transparência e explicabilidade dos sistemas.

Quando algoritmos operam sem clareza sobre seus critérios de decisão, podem gerar impactos negativos significativos, como injustiças, discriminações e falhas operacionais, de modo que, a mitigação desses riscos exige uma abordagem ética *by design* nos projetos, desde seu o desenvolvimento até a aplicação efetiva da inteligência artificial em rotinas, com atenção especial à governança e ao compliance digital.

Algoritmos de inteligência artificial podem ser contaminados por vieses (bias) originados nos valores pessoais dos próprios desenvolvedores ou influenciados por interesses de stakeholders envolvidos em sua evolução. Esses vieses, muitas vezes inconscientes, podem se infiltrar nos dados utilizados para treinar os modelos ou nas decisões tomadas durante o design do sistema.

O resultado é um algoritmo que reflete preferências, preconceitos ou objetivos específicos, comprometendo sua imparcialidade e equidade, ou seja, "*tanto a construção da base de dados usada pela IA, quanto as fórmulas matemáticas aplicadas a estes dados em vistas a resultados – tais como previsão ou recomendação – podem introduzir preconceitos ou inadequações que fazem parte da estrutura dos sistemas de AI*" (MULHOLAND – 2022)

Além disso, a opacidade das máquinas, ou seja, a dificuldade de compreender como a inteligência artificial chegou à determinada conclusão, pode gerar consequências graves, entre elas, destaca-se a discriminação de clientes em processos de elegibilidade para serviços financeiros, seguros ou benefícios.

Como exemplo concreto disso, temos o uso de marketing segmentado de forma negativa, explorando vulnerabilidades ou

reforçando estereótipos, bem como sistemas de recrutamento baseados em inteligência artificial, que podem excluir candidatos por critérios etaristas, racistas ou misóginos, mesmo que isso não seja intencional.

Outro fator crítico é a qualidade dos bancos de dados utilizados para treinar os sistemas de inteligência artificial, uma vez que, quando esses dados são incompletos, desatualizados ou enviesados, os resultados esperados não são atingidos, de modo que, o sistema pode tomar decisões equivocadas.

Nesse contexto, a norma ISO/IEC 24368:2022 surge como uma diretriz essencial, pois orienta sobre as melhores práticas que preservam os sistemas de inteligência artificial contra os impactos negativos de algoritmos inadequados, promovendo uma abordagem ética desde a origem dos dados até a aplicação dos modelos. A norma reforça ainda que a ética algorítmica deve ser incorporada à governança corporativa, tornando-se uma responsabilidade direta do compliance digital.

A norma ressalta então a necessidade de tratamento eficiente de banco de dados utilizados na inteligência artificial, ou seja, desde a sua coleta, preparação e armazenamento seguro dos dados, bem como o controle de qualidade, que são etapas fundamentais para garantir que a inteligência artificial funcione de forma ética e eficaz, conforme o seu item 4.2:

> **4.2 Fundamental sources**
>
> Without data, the development and use of AI cannot be possible. Therefore, the importance of data and data quality makes traceability and data management a pivotal consideration in the use and development of AI. The following data-oriented elements are at the core of creating ethical and sustainable AI:
>
> — data collection (including the means or measures of such data collection);
>
> — data preparation;
>
> — monitoring of traceability;
>
> — access and sharing control (authentication);
>
> — data protection;
>
> — storage control (adding, change, removal);
>
> — data quality.

Não bastasse isso, para que os sistemas operem sem vieses, é fundamental ainda a utilização de frameworks éticos baseados em outros normativos, como a Declaração Universal de Direitos Humanos, para entendimento e aplicação de valores como honestidade, respeito etc.

Além disso, a norma destaca a importância da *accountability,* ou seja, da responsabilidade que a empresa deve assumir diante de eventuais questões éticas que possam surgir com a evolução dos sistemas de inteligência artificial.

Por fim, como esses sistemas estão em constante transformação, é inevitável que surjam situações imprevisíveis, pelo que, cabe à organização definir ações apropriadas, estabelecer mecanismos de resposta rápida e garantir que a ética continue sendo um pilar central em sua estratégia tecnológica.

12.5. POLÍTICAS DE SEGURANÇA DA INFORMAÇÃO PARA GOVERNANÇA DE INTELIGÊNCIA ARTIFICIAL

Com o crescimento exponencial da inteligência artificial nas operações empresariais, torna-se imperativo estabelecer também políticas específicas de segurança da informação em seu funcionamento, as quais garantam sua proteção, eficiência e conformidade, o que possibilitará sua melhor governança, senão vejamos:

Os sistemas de inteligência artificial, pelo fato de utilizarem grandes volumes de dados, exigem uma proteção ainda mais robusta do que os sistemas convencionais, sendo que, um dos aspectos mais desafiadores é sua capacidade de mudar de comportamento ao longo do tempo, especialmente quando integrada a processos dinâmicos.

Essa característica exige que as empresas estejam preparadas para realizar adaptações periódicas em seus sistemas, ajustando parâmetros, atualizando modelos e revisando políticas conforme o contexto operacional evolui. A gestão de inteligência artificial, portanto, não pode ser estática, ela deve ser flexível, responsiva e continuamente monitorada.

Essa gestão diferenciada tem como principal objetivo proteger os sistemas de inteligência artificial contra ataques cibernéticos e falhas operacionais, de forma que, a sua proteção deverá envolver

camadas adicionais de segurança, com foco em integridade, confidencialidade e disponibilidade dos dados e modelos.

Logo, para garantir essa proteção, é essencial implementar políticas de segurança da informação específicas para inteligência artificial, as quais devem contemplar não apenas os aspectos técnicos, mas também os recursos humanos e organizacionais necessários para sua execução.

Tais políticas incluem a capacitação contínua de profissionais envolvidos, a alocação de recursos adequados e a definição clara de responsabilidades, pelo que, a gestão eficiente da inteligência artificial depende de uma abordagem integrada que envolva tecnologia, pessoas e processos, da mesma forma que os sistemas de informação convencionais.

Nesse contexto, a norma ISO/IEC 42001 surge como referência fundamental, pois estabelece alguns princípios orientativos para a gestão de sistemas de inteligência artificial, oferecendo diretrizes para o desenvolvimento de políticas específicas que assegurem a segurança e a conformidade desses sistemas, conforme trecho de sua introdução:

> The Al management system should be integrated with the organization's processes and overall management structure. Specific issues related to Al should be considered in the design of processes information systemsand controls. Crucial etamples of such management processes are:
>
> - determination of organizational objectives, involvement of interested parties and organizational policy;
>
> - management of risks and opporrunities;
>
> -processes for the management of concerns related to the trustworthiness of Al systems such as security, safety, fairness transparency, data quality and quality of Al systems throughout their life cycle;
>
> - processes for the management of suppliers, partners and third parties that provide or develop Al systems for the organization.
>
> This document provides guidelines for the deployment of applicable controls to support such processes.

A norma propõe, então, diversas medidas, como a criação de evidências concretas da existência e aplicação destas políticas

especiais, o que significa que as organizações devem documentar seus processos, registrar decisões, manter logs de auditoria e demonstrar, de forma transparente, que suas práticas estão alinhadas com os princípios regulatórios aplicáveis, pois a rastreabilidade e a prestação de contas tornam-se elementos essenciais da governança algorítmica.

Além disso, é necessário desenvolver processos específicos que se integrem aos sistemas já existentes, como a gestão de privacidade e proteção de dados, tendo em vista que, a sinergia entre essas áreas permite uma abordagem mais completa e eficaz, garantindo que os sistemas de inteligência artificial respeitem os direitos dos titulares de dados e estejam em conformidade com legislações como a LGPD e o GDPR.

Além disso, as políticas de gestão de inteligência artificial também devem ser estendidas aos fornecedores e parceiros comerciais, pela inclusão de cláusulas contratuais específicas sobre segurança, conformidade e responsabilidade no uso desta tecnologia, como medida preventiva importante para proteger os sistemas corporativos.

A gestão de terceiros torna-se, assim, uma extensão da governança interna, com impacto direto na segurança e na integridade dos sistemas.

A norma também destaca como outro ponto crítico o fator humano, tendo em vista que, os profissionais envolvidos no desenvolvimento, operação e supervisão da inteligência artificial representam um vetor de risco que deve ser considerado nas políticas de gestão.

Ora, a negligência, o desconhecimento ou mesmo a má-fé podem resultar em vazamentos de dados, falhas operacionais ou vulnerabilidades exploráveis, de modo que, a capacitação, a conscientização e o monitoramento contínuo são estratégias fundamentais para mitigar esses riscos.

Determina-se ainda que, tais políticas sejam disseminadas de forma clara e acessível a todos os usuários dos sistemas de inteligência artificial com a criação de uma cultura corporativa voltada para o uso controlado e responsável da inteligência artificial evitando-se práticas informais e não autorizadas, como a chamada *shadow AI*,

que pode comprometer toda a arquitetura de segurança cibernética da empresa.

Vale destacar que, a questão da *shadow AI,* ou seja, o uso de ferramentas de inteligência artificial não autorizada pelas empresas, infelizmente tornou-se um problema bastante grave e comum, porque sistemas de informação são projetados para estarem em funcionamento coordenado com alguns tipos de inteligência artificial escolhidos por gestores, pelo que, não se pode permitir que alguma outra tecnologia desautorizada (ex: Chatgpt, Copilot, Gemini e etc.) seja instalada por colaboradores, prejudicando a segurança cibernética corporativa.

Além disso, as políticas de gestão de inteligência artificial devem estar alinhadas com outras políticas organizacionais, como as de compliance, governança de dados, privacidade, segurança da informação, ou seja, a integração dessas políticas garante consistência, evita conflitos normativos e fortalece a estrutura de controle interno da empresa.

Outro ponto fundamental é que, ao definir políticas de segurança da informação para inteligência artificial, sejam estabelecidos claramente os níveis de risco associados a cada sistema e os modos de tratamento desses riscos, efetuando-se a sua classificação quanto à criticidade, a definição de planos de contingência, a realização de testes periódicos e a revisão contínua das políticas.

CONCLUSÕES

Podemos então concluir que é fundamental para a eficiente governança de inteligência artificial que sejam adotadas as orientações dos standards de Compliance Digital de diversas normas ISO e NIST, os quais trazem orientações essenciais para a gestão segura destas soluções, aplicando-se boas práticas em seus processos, de forma a evitar-se problemas relacionados à transparência de algoritmos, ética e resultados ruins em sua performance.

REFERÊNCIAS

DEEP LEARNING. Disponível em: https://www.salesforce.com/br/blog/machine=-learning-vs-deep-learning/?d701=e00000dDLNgOAAX&nc-701ed00000DLR7FAAX&utm_source=google&utm_medium=paid_sear-

ch&utm_campaign=latam_br_all-ai-and-data-apm-l2_cross-industry&utm_content=all-segments_pg-pt-mash_701ed00000DLNgOAAX_portuguesebr_machine-learning-vs-deep-learning&utm_term=machinelearning&ef_id=CjwKCAjw3tzHBhBREiwAlMJoUn2FcS7ZysS_oxDYHR--keDPApOmJ-cBX8S4RYaRkeYNvbjn-tNf01RoCn24QAvD_BwE:G:s&gclsrc=aw.ds&&pcrid=772823545571&pdv=c&gad_source=1&gad_campaignid=22992518341&gbraid=0AAAAAD8tsjlsREKWJ3m2xzJIE7EBe41kb&gclid=CjwKCAjw3tzHBhBREiwAlMJoUn2FcS7ZysS_oxDYHR-keDPApOmJ-cBX8S4RYaRkeYNvbjn-Nf01RoCn24QAvD_BwE. Acesso em 21.10.2025.

ISO (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION). **2020**. *ISO 24028:2020 Overview of trustworthiness in artificial intelligence— Requirements*. 1st ed. Geneva: ISO, 2020. Disponível em: https://www.iso.org/standard/24028.pdf. Acesso em: 10.10.2025.

ISO (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION). **2022**. *ISO 38507:2022 Governance implications of the use of artificial intelligence by organizations— Requirements*. 1st ed. Geneva: ISO, 2022. Disponível em: https://www.iso.org/standard/38507.pdf. Acesso em: 10.10.2025.

ISO (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION). ISO 24368:2022 — Overview of ethical and societal concerns. Geneva: ISO, 2022. Disponível em: https://www.iso.org/standard/78507.html. Acesso em: 10.10.2025.

ISO (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION). ISO/IEC 42001:2023 — Artificial intelligence — Management system. Geneva: ISO, 2023. Disponível em: https://www.iso.org/standard/81230.html. Acesso em: 10.10.2025.

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY. **NIST Artificial Intelligence Risk Management**. 2. ed., Washington, DC, National Institute of Standards and Technology, 2023. Disponível em: https://www.nist.gov/cyberframework Acesso em: 15.set. 2025. Acesso em: 10.10.2025.

MACHINE LEARNING. Disponível em: https://www.salesforce.com/br/blog/machine=-learning-vs-deep-learning/?d701=e00000dDLNgOAAX&nc-701ed00000DLR7FAAX&utm_source=google&utm_medium=paid_search&utm_campaign=latam_br_all-ai-and-data-apm-l2_cross-industry&utm_content=all-segments_pg-pt-mash_701ed00000DLNgOAAX_portuguesebr_machine-learning-vs-deep-learning&utm_term=machinelearning&ef_id=CjwKCAjw3tzHBhBREiwAlMJoUn2FcS7ZysS_oxDYHR-keDPApOmJ--cBX8S4RYaRkeYNvbjn-tNf01RoCn24QAvD_BwE:G:s&gclsrc=aw.ds&&pcrid=772823545571&pdv=c&gad_source=1&gad_campaignid=22992518341&gbraid=0AAAAAD8tsjlsREKWJ3m2xzJIE7EBe41kb&gclid=CjwKCAjw3tzHBhBREiwAlMJoUn2FcS7ZysS_oxDYHR-keDPApOmJ-cBX8S4RYaRkeYNvbjn--Nf01RoCn24QAvD_BwE. Acesso em 17.10.2025.

MULHOLAND, Caitlin, Direitos Fundamentais e Sociedade Tecnológica, org. CARPENA, Heloísa, Martins, Guilherme Magalhães e SCHREIBER, Anderson, São Paulo, Editora Foco, 1ª ed., 2022, p. 172.

REASONING. Disponível em: https://www.ibm.com/think/topics/ai-reasoning. Acesso em 21.10.2025.