PROVVI

Tecnologia de Autenticidade de Imagens

A Última Milha da Auditabilidade:

C2PA e ICP-Brasil como resposta ao PL 2338/2023

Marco Legal da IA e a Prova de Autenticidade que a Responsabilidade Objetiva vai Exigir

Produzido pelo Comitê de Tecnologia e Compliance da Provvi · Março de 2026

O PROBLEMA CENTRAL

Sistemas de IA que decidem sobre sinistros, crédito e seguros alimentam-se de imagens digitais. O PL 2338/2023 exige auditabilidade das decisões — mas a autenticidade dos dados de entrada nunca é verificada.

A SOLUÇÃO PROPOSTA

O padrão C2PA com assinatura ICP-Brasil estabelece uma cadeia de custódia criptográfica da captura da imagem até a decisão do modelo — tornando o input do sistema de IA tão auditável quanto o output.

provvi.com.br · contato@provvi.com.br

Sumário Executivo

O Marco Legal da Inteligência Artificial — PL 2338/2023, aprovado pelo Senado em dezembro de 2024 e em análise pela Câmara dos Deputados — estabelece um princípio que vai transformar a operação de seguradoras, fintechs e qualquer empresa que use IA em processos decisórios: a responsabilidade objetiva por danos causados por sistemas de alto risco.

Responsabilidade objetiva significa que, em caso de dano, a empresa responde independentemente da comprovação de culpa. Para sistemas de alto risco — e sistemas de IA que decidem sobre sinistros, crédito e apólices se enquadram nessa categoria —, a carga probatória inverte: é a empresa que precisa provar que seu sistema funcionou corretamente, não o consumidor que precisa provar que foi prejudicado.

Essa inversão cria uma exigência técnica que o setor ainda não discutiu com a profundidade que merece: para provar que um sistema de IA funcionou corretamente, é necessário provar que os dados que alimentaram a decisão eram autênticos. Uma decisão de IA é tão auditável quanto as evidências que a sustentam.

Este whitepaper apresenta três argumentos:

- O PL 2338/2023 cria um regime de responsabilidade que exige auditabilidade de ponta a ponta — da evidência coletada à decisão tomada. A maioria dos sistemas atuais é auditável na camada do modelo, mas não na camada da evidência.
- C2PA com ICP-Brasil resolve a 'última milha' da auditabilidade: a autenticidade da imagem ou documento que serve de input para o sistema de IA.
- A implementação antecipada — antes da aprovação definitiva da lei — é uma vantagem de compliance que reduz risco jurídico e eleva a qualidade probatória em processos administrativos e judiciais.

O argumento não é que C2PA resolve todos os requisitos do Marco Legal da IA. É que sem autenticidade verificável de evidências, nenhum sistema de IA — por mais sofisticado que seja — consegue demonstrar que sua cadeia de raciocínio partiu de dados confiáveis. Essa é a lacuna que o Provvi endereça.

1. O Marco Legal da IA: O Que o PL 2338/2023 Realmente Implica

1.1 Estado Atual da Tramitação

O PL 2338/2023, de autoria do Senador Rodrigo Pacheco, foi aprovado pelo Senado Federal em 10 de dezembro de 2024 e segue em análise pela Câmara dos Deputados, sob relatoria do deputado Aguinaldo Ribeiro (PP-PB). O texto ainda deve retornar ao Senado para análise das emendas da Câmara antes de seguir para sanção presidencial.

A votação, inicialmente prevista para o final de 2024, foi adiada após impasses em torno de pontos sensíveis — direitos autorais no uso de obras em treinamento de IA e a definição precisa das exceções para sistemas de alto risco. A expectativa do setor é de aprovação ao longo de 2026, com vigência progressiva após a sanção.

Paralelamente, o Poder Executivo encaminhou ao Congresso um projeto complementar que cria o Sistema Nacional para Desenvolvimento, Regulação e Governança de Inteligência Artificial (SIA) e estabelece a Autoridade Nacional de Proteção de Dados (ANPD) como autoridade regulatória central — corrigindo um vício de iniciativa identificado no texto original do Senado.

> **Ponto de atenção:** O texto ainda tramita, mas suas linhas estruturais estão consolidadas desde a aprovação no Senado. As empresas que aguardam a sanção presidencial para iniciar adequação técnica estarão operando em déficit de compliance desde o primeiro dia de vigência.

1.2 A Estrutura de Risco do PL 2338/2023

A espinha dorsal do PL é a classificação de sistemas de IA por nível de risco, inspirada no AI Act da União Europeia. O projeto brasileiro define duas categorias principais:

Risco Excessivo — Uso Proibido

Sistemas que exploram vulnerabilidades de pessoas, manipulam comportamento de forma subliminar, aplicam pontuação social injusta em contextos governamentais, ou operam como armas autônomas. Esses usos são vedados independentemente de qualquer justificativa.

Alto Risco — Sujeito a Governança Estrita

Sistemas com impacto direto em direitos fundamentais: saúde, justiça, segurança, crédito, emprego, educação, infraestrutura crítica e identificação biométrica. Para esse grupo, o PL impõe obrigações específicas de governança, transparência e responsabilidade.

Para o setor de seguros e serviços financeiros, a classificação de risco é direta: sistemas que determinam a aceitação de uma apólice, o valor de um prêmio, a aprovação ou negação de um sinistro, ou a detecção de fraude são sistemas que influenciam decisões com impacto direto nos direitos e no patrimônio do segurado. A enquadramento como Alto Risco não é uma interpretação — é a consequência natural do texto.

1.3 As Obrigações para Sistemas de Alto Risco

Para sistemas classificados como Alto Risco, o PL 2338/2023 estabelece um conjunto de obrigações que impactam diretamente a arquitetura técnica das soluções:

Obrigação	Implicação Técnica
Avaliação de Impacto Algorítmico (AIA)	Documentação formal dos riscos de viés, falhas e danos potenciais do sistema antes da operação.

Transparência e explicabilidade	Capacidade de explicar para o afetado por que o sistema tomou determinada decisão.
Supervisão humana contínua	Processos de revisão humana em decisões de alto impacto — especialmente negações e sanções.
Responsabilidade civil objetiva	A empresa responde por danos independentemente de culpa; ônus da prova inverte para o operador.
Rastreabilidade da decisão	Capacidade de reconstruir o raciocínio do sistema — incluindo os dados usados como input.
Auditabilidade por autoridade regulatória	A ANPD poderá requisitar acesso a logs, dados, modelos e evidências para fins de fiscalização.

2. O Problema da Última Milha: A Evidência que o Modelo Não Verifica

2.1 A Cadeia de Decisão de um Sistema de IA em Seguros

Um sistema de IA que avalia um sinistro de automóvel opera sobre uma cadeia de entradas: fotos do veículo danificado, dados do histórico do segurado, informações do boletim de ocorrência, resultado de vistoria prévia, laudos técnicos. O modelo processa essas entradas e produz uma saída — uma recomendação de aprovação, negação ou investigação.

A auditabilidade que o PL 2338/2023 exige concentra-se, nos debates atuais, na camada do modelo: como ele tomou a decisão, quais variáveis pesou, se há viés sistemático, se a explicação oferecida ao segurado é compreensível. Essa auditabilidade é necessária — e já existem ferramentas técnicas razoavelmente maduras para implementá-la.

O que não está sendo discutido com a mesma profundidade é a auditabilidade da camada anterior: a autenticidade das evidências que alimentaram o modelo. E é exatamente aqui que a vulnerabilidade jurídica mais significativa reside.

> **O argumento central:** Uma cadeia de raciocínio auditável que parte de evidências não verificáveis não é realmente auditável. Se o input do modelo pode ser manipulado sem detecção, a auditabilidade do modelo é uma garantia incompleta.

2.2 Os Vetores de Comprometimento de Evidência

No contexto de seguros, as evidências digitais que alimentam sistemas de IA são primariamente imagens — fotos de sinistro, laudos de vistoria, registros de dano. Esses arquivos percorrem um caminho entre a captura original e a entrada no sistema de análise. Em cada ponto desse caminho, há uma oportunidade de comprometimento:

Substituição de imagem pelo segurado

O segurado fotografa um veículo com danos maiores do que os do seu próprio carro, ou adiciona danos digitalmente à imagem antes de submeter ao processo de sinistro. Sem verificação de autenticidade na captura, a imagem submetida é indistinguível da original para o sistema de IA.

Reutilização de imagem legítima

Uma imagem autêntica de um sinistro anterior — do mesmo veículo ou de outro — é submetida como evidência de um novo evento. A imagem passa em qualquer verificação de integridade baseada em hash do arquivo, mas não corresponde ao evento declarado.

Comprometimento do canal de transmissão

A imagem é capturada legitimamente, mas substituída durante o upload ao sistema da seguradora ou da empresa de gestão de sinistros. Sem prova criptográfica de que a imagem não foi alterada entre a captura e o recebimento, a substituição é indetectável.

Fraude organizada com banco de imagens

Grupos organizados de fraude operam com bibliotecas de imagens de danos legítimos, selecionando a imagem mais adequada para cada perfil de veículo e tipo de sinistro declarado. Sistemas de detecção de fraude baseados em análise de imagem são inúteis se a imagem em si é tecnicamente íntegra mas contextualmente falsa.

2.3 A Responsabilidade Objetiva e o Problema da Evidência

O regime de responsabilidade objetiva do PL 2338/2023 cria um cenário específico para seguradoras: se um sistema de IA nega indevidamente um sinistro legítimo, a empresa responde pelo dano. Mas há um corolário menos discutido: se o sistema aprova um sinistro fraudulento porque a evidência submetida era falsa, a empresa sofre o prejuízo financeiro e ainda pode ser questionada sobre a qualidade de seus controles de auditabilidade.

Em ambos os casos — falso negativo e falso positivo —, a defesa jurídica da seguradora requer a capacidade de demonstrar que o sistema de IA operou sobre evidências verificáveis. Sem essa capacidade, qualquer contestação administrativa ou judicial é substancialmente mais difícil.

Cenário	Risco sem Autenticidade C2PA	Posição com Autenticidade C2PA
Sinistro legítimo negado indevidamente	Empresa não consegue provar que a decisão partiu de evidência correta	Empresa demonstra que a evidência era autêntica; o problema está no modelo — passível de correção
Sinistro fraudulento aprovado	Empresa não consegue provar que a evidência foi adulterada antes de entrar no sistema	Empresa demonstra adulteração pós-captura; responsabilidade desloca-se para o fraudador
Auditoria regulatória da ANPD	Logs do modelo disponíveis, mas autenticidade das evidências não verificável	Cadeia completa auditável: da captura ao manifesto à decisão do modelo
Processo judicial do segurado	Empresa não tem como provar que a imagem que o modelo viu é a imagem que o segurado enviou	Empresa apresenta manifesto C2PA como evidência forense — admissível e verificável por qualquer perito

3. C2PA como Camada de Auditabilidade de Evidência

3.1 O que C2PA Resolve neste Contexto

C2PA (Coalition for Content Authenticity and Provenance) é um padrão técnico aberto que define como metadados de proveniência devem ser estruturados, assinados e verificados em arquivos de mídia. Desenvolvido por um consórcio que inclui Adobe, Microsoft, Google, Intel, Reuters, BBC e mais de 300 organizações, ele foi originalmente concebido para combater a chamada Fraude Sintética — deepfakes e imagens geradas por IA em contexto jornalístico — mas sua aplicação natural e mais impactante no setor de seguros é outra: a Fraude de Contexto. A distinção é fundamental. Fraude Sintética é a criação de uma imagem falsa — um veículo danificado que não existe, gerado por IA generativa. É o tipo de fraude que domina o debate público sobre deepfakes, mas que ainda representa uma fração do problema no setor de seguros. Fraude de Contexto é o uso de uma imagem completamente real em um contexto falso: a foto autêntica de um carro batido submetida como evidência de um sinistro diferente; a imagem legítima de uma colisão de 2023 usada para registrar um evento de 2026; o veículo de outra pessoa fotografado para fraudar uma apólice própria. A imagem passa em qualquer análise de integridade técnica — porque ela é tecnicamente íntegra. O problema não é o conteúdo da foto, é a mentira sobre quando, onde e por quem ela foi capturada. C2PA endereça diretamente a Fraude de Contexto, que é, segundo dados do mercado segurador, a modalidade dominante e mais custosa. Ao vincular criptograficamente a imagem ao dispositivo de captura, ao local GPS e ao timestamp do momento real da captura, C2PA torna impossível descontextualizar uma imagem autêntica sem invalidar sua prova de proveniência.

No contexto do PL 2338/2023 e de sistemas de IA em seguros, C2PA resolve especificamente a lacuna de auditabilidade de evidência: a capacidade de provar que a imagem que entrou no sistema de IA é a mesma imagem capturada no momento e local declarados, sem modificações.

Um manifesto C2PA é gerado no momento da captura e contém:

- Identidade criptográfica do dispositivo de captura — hardware e aplicativo.
- Hash SHA-256 do frame original, calculado antes de qualquer compressão ou processamento.
- Timestamp criptograficamente vinculado ao momento da captura — não ao momento do upload.
- Localização GPS integrada ao manifesto — não ao sistema receptor.
- Asserções de negócio customizáveis — tipo de evento, identificação do caso, dados do operador.
- Cadeia de assinaturas verificável por qualquer parte, sem necessidade de acesso ao sistema de origem.

Qualquer alteração no arquivo após a captura — substituição de pixel, modificação de metadado, recorte ou ajuste de cor — invalida a assinatura e torna a adulteração detectável. Isso é verificável por qualquer ferramenta compatível com o padrão, incluindo verificadores públicos.

3.2 C2PA + ICP-Brasil: Validade Jurídica no Contexto Brasileiro

Para que a autenticidade verificada pelo padrão C2PA tenha validade jurídica plena no Brasil — e não apenas integridade técnica —, a assinatura deve ser emitida por uma autoridade certificadora reconhecida pela ICP-Brasil, a Infraestrutura de Chaves Públicas Brasileira regulamentada pela MP 2.200-2/2001 e reconhecida como equivalente à assinatura de próprio punho pela legislação nacional.

A arquitetura Provvi implementa dois estágios de assinatura:

Etapa	Local	O que garante

1	Dispositivo móvel no momento da captura	Hash SHA-256, GPS, timestamp e identidade do dispositivo vinculados ao arquivo — prova de integridade da captura antes de qualquer transmissão.
2	Backend seguro (AWS KMS + cert. ICP-Brasil)	Re-assinatura com certificado ICP-Brasil via ECDSA P-256 — validade jurídica equivalente à assinatura de próprio punho, reconhecida pela legislação brasileira.
Resultado	Manifesto duplamente assinado	Evidência com cadeia de custódia completa: integridade técnica + validade jurídica + verificação pública independente.

3.3 Posicionamento na Cadeia de Auditabilidade do Sistema de IA

C2PA não substitui as ferramentas de explicabilidade e auditabilidade do modelo de IA — XAI (Explainable AI), logs de decisão, análises de viés. Ele complementa essas ferramentas ao fechar a lacuna que elas deixam em aberto: a autenticidade da evidência de entrada.

Camada	O que audita	Ferramentas existentes
Evidência (input)	Autenticidade e proveniência da imagem ou documento que alimentou o modelo	C2PA + ICP-Brasil — lacuna atual na maioria dos sistemas
Modelo (processamento)	Pesos, parâmetros, viés, explicabilidade da decisão	XAI, SHAP, LIME, auditoria de modelo
Decisão (output)	Log da decisão, justificativa oferecida ao afetado, histórico de revisão humana	Sistemas de log, workflow de revisão, CRM de sinistros

> **Nota:** A linha destacada em âmbar representa a camada ainda não coberta na maioria dos sistemas de IA em seguros. C2PA preenche especificamente essa lacuna, sem requerer modificação das camadas de modelo e decisão já implementadas.

4. Aplicação Prática: Casos de Uso no Setor de Seguros

4.1 Sinistro de Automóvel

O processo de abertura de sinistro de automóvel é o caso de uso mais direto. O segurado registra o evento e envia fotos do dano. Essas imagens são o input primário do sistema de IA que classifica o sinistro, estima o valor do dano e recomenda aprovação, negação ou investigação.

Com C2PA integrado ao aplicativo de registro de sinistro:

1. O segurado abre o app e aciona a câmera pelo SDK Provvi integrado.
2. Cada foto capturada recebe um manifesto C2PA imediato — hash, GPS, timestamp, identidade do dispositivo.
3. As imagens são transmitidas ao backend da seguradora com o manifesto anexado.
4. O sistema de IA recebe as imagens junto com a confirmação de autenticidade C2PA.
5. A decisão do modelo é logada com referência ao session_id das imagens autenticadas.
6. Em caso de contestação, a cadeia completa é verificável: da captura à decisão.

O diferencial para o compliance do PL 2338/2023: a seguradora consegue demonstrar que a rastreabilidade da decisão alcança a camada de evidência, não apenas a camada do modelo.

4.2 Detecção de Fraude em Sinistro

Sistemas de IA de detecção de fraude analisam padrões em imagens — consistência de iluminação, metadados inconsistentes, sinais de edição digital. Esses sistemas são eficazes contra fraudes técnicas, mas têm uma vulnerabilidade fundamental: não conseguem detectar uma imagem que é tecnicamente íntegra mas contextualmente falsa (imagem de outro veículo, de outro evento, de outra data).

C2PA resolve essa vulnerabilidade complementando a análise do modelo com verificação de proveniência. Uma imagem com manifesto C2PA válido prova não apenas que não foi editada digitalmente, mas que foi capturada naquele dispositivo, naquele local, naquele momento. Uma imagem sem manifesto válido — ou com manifesto inválido — é um sinal de alerta independente da análise de conteúdo.

> **Benefício operacional:** A combinação de análise de conteúdo (modelo de IA) com verificação de proveniência (C2PA) cria dois vetores independentes de detecção de fraude. Uma fraude sofisticada que engana o modelo de imagem ainda precisa forjar um manifesto criptográfico válido — o que é computacionalmente inviável sem acesso à chave privada do dispositivo.

4.3 Vistoria Prévia para Contratação de Seguro

A vistoria prévia — exigida em apólices de veículo como condição de cobertura — é um ponto crítico de risco de fraude na subscrição. O segurado pode submeter imagens de um veículo em melhor estado do que o real, ou ocultar pré-existências que excluiriam a cobertura ou elevariam o prêmio.

Com C2PA na captura da vistoria prévia, a seguradora possui uma evidência criptograficamente verificável do estado do veículo no momento da contratação. Em caso de sinistro, a comparação entre o estado documentado na vistoria prévia e o estado declarado no sinistro é auditável de forma objetiva — eliminando disputas sobre a condição original do bem segurado.

4.4 Avaliação de Risco com Dados de Terceiros

Seguradoras cada vez mais integram dados de fontes externas em seus modelos de avaliação de risco — laudos de vistoria de ECVs, relatórios de inspeção, documentação de histórico do veículo. Quando esses dados chegam de terceiros, a questão de autenticidade é ainda mais relevante: a seguradora não controla o processo de captura.

A integração do SDK Provvi pelas ECVs e empresas de vistoria parceiras cria um ecossistema onde os laudos chegam à seguradora com autenticidade já verificada na origem. O fluxo de dados de terceiros passa a ter a mesma rastreabilidade criptográfica que os dados capturados diretamente pelo segurado.

5. Implementação: Integração na Arquitetura Existente

5.1 Princípio de Integração

A integração de C2PA em sistemas de IA de seguros não requer substituição de infraestrutura existente. O princípio é de adição de camada: o SDK Provvi é inserido no ponto de captura de imagem — o aplicativo do segurado, o app de vistoria, a ferramenta do regulador de sinistros —, e o manifesto gerado é transmitido junto com o arquivo de imagem para os sistemas já existentes.

Os sistemas de IA, CRM de sinistros e plataformas de gestão não precisam ser modificados para receber imagens com manifesto C2PA. A verificação de autenticidade pode ser feita em uma camada separada, antes ou durante a entrada dos dados no modelo, sem alteração do pipeline de decisão existente.

5.2 Componentes Técnicos

- SDK Android e iOS — integrado ao aplicativo da seguradora ou do prestador de serviço; substitui a câmera nativa pelo pipeline de captura autenticada.
- API de verificação — endpoint que recebe o session_id ou o hash da imagem e retorna o status de autenticidade, os metadados de proveniência e o resultado da verificação de assinatura ICP-Brasil.
- Backend de assinatura — serviço em nuvem que aplica a segunda assinatura com certificado ICP-Brasil; operado pelo Provvi como serviço ou integrado à infraestrutura própria da seguradora.
- Log de auditoria — registro imutável do manifesto, associado ao processo de sinistro ou contratação no sistema de gestão da seguradora.

5.3 Fluxo de Integração com Sistema de IA Existente

Cadeia de custódia da evidência — do smartphone ao motor de IA

Integração plug-and-play: o SDK Provvi é inserido no ponto de captura; os sistemas existentes de IA e CRM não requerem modificação.

#	Etapa	Detalhe técnico
1	**Captura via SDK Provvi**	App chama SDK em vez de câmera nativa; manifesto C2PA gerado automaticamente na captura.
2	**Transmissão com manifesto**	Imagem + manifesto transmitidos ao backend; hash local garante integridade durante transmissão.
3	**Assinatura ICP-Brasil no backend**	Backend Provvi aplica segunda assinatura via AWS KMS com certificado ICP-Brasil; session_id gerado.
4	**Verificação antes do modelo**	Sistema da seguradora consulta API Provvi com session_id; recebe confirmação de autenticidade antes de encaminhar ao modelo de IA.
5	**Input no modelo de IA**	Imagem autenticada entra no pipeline de IA; status de autenticidade é logado junto com os demais metadados do processo.

6	**Log de auditoria**	Session_id e resultado de verificação são associados ao processo de sinistro no CRM — cadeia de custódia completa disponível para auditoria.

5.4 Impacto no Desempenho Operacional

A adição da camada C2PA ao processo de captura introduz latência mínima — o manifesto é gerado no dispositivo em tempo real, sem necessidade de roundtrip de rede. A assinatura ICP-Brasil no backend ocorre de forma assíncrona: o usuário recebe o retorno da captura em cerca de 3,5 segundos, e a assinatura backend é aplicada em background sem impactar a experiência de uso.

Para processos offline — captura em locais com conectividade limitada — o SDK mantém fila local: a assinatura backend é aplicada quando a conectividade é restaurada, sem comprometer a integridade do manifesto gerado na captura.

6. O Argumento Regulatório: Compliance Proativo

6.1 Sanções do PL 2338/2023

O regime sancionatório previsto no PL 2338/2023 para sistemas de alto risco é substancial: multas de até R$ 50 milhões ou 2% do faturamento do grupo econômico infrator, além de advertências, suspensão ou proibição definitiva do uso do sistema. A ANPD terá poder de requisitar acesso a documentação técnica, logs e dados de treinamento.

Em um processo de investigação regulatória, a capacidade de demonstrar que o sistema de IA operou sobre evidências verificáveis é um diferencial significativo. A ausência dessa demonstração não é apenas uma lacuna técnica — é um argumento favorável ao enquadramento como omissão negligente na implementação de controles adequados.

6.2 A Relação com LGPD e ANPD

A ANPD, designada como autoridade regulatória central do SIA, já possui competência consolidada em proteção de dados pessoais e tende a interpretar as obrigações do Marco Legal da IA com um olhar rigoroso de qualidade e integridade de dados. Há aqui uma conexão direta com a LGPD que merece atenção específica: o Art. 6º da Lei 13.709/2018 estabelece, entre os princípios que regem o tratamento de dados pessoais, o princípio da Exatidão — a obrigação de que os dados sejam exatos, claros, relevantes e atualizados conforme a necessidade do tratamento. Um sistema de IA de alto risco que toma decisões sobre sinistros, crédito ou apólices alimentando-se de imagens não autenticadas está, na prática, operando sobre dados cuja exatidão não pode ser verificada — e portanto sobre dados que podem não ser exatos. Não se trata apenas de uma falha técnica de segurança: é uma potencial violação direta do princípio da Exatidão da LGPD. A ANPD tem fundamento legal para enquadrar o uso de evidências não verificáveis em sistemas de IA de alto risco como descumprimento simultâneo da LGPD e do Marco Legal da IA — duplicando a base de responsabilização disponível ao regulador.

A implementação de C2PA deve observar os princípios da LGPD quanto aos dados capturados no manifesto — GPS, identidade do dispositivo, timestamp. A base legal para o tratamento é o cumprimento de obrigação legal e o interesse legítimo do operador, e a minimização de dados deve ser aplicada: o manifesto registra apenas os dados necessários para proveniência.

6.3 Precedentes Internacionais

A convergência entre autenticidade de evidências digitais e responsabilidade de sistemas de IA não é uma preocupação exclusivamente brasileira. O AI Act europeu, em vigor desde agosto de 2024, estabelece obrigações similares de rastreabilidade de dados para sistemas de alto risco — e a interpretação regulatória europeia tem sido de que a cadeia de auditabilidade deve alcançar os dados de entrada, não apenas o modelo.

A indústria de seguros europeia, sob pressão do AI Act, tem adotado padrões de proveniência de evidência que antecipam os requisitos regulatórios. O Brasil, com o PL 2338/2023 alinhado ao mesmo modelo de risco do AI Act, seguirá trajetória similar. Empresas que implementam agora não apenas cumprem antecipadamente — constroem vantagem competitiva em um mercado que vai exigir essa capacidade.

7. Conclusão: Governança de IA e a Proteção do Patrimônio Corporativo

O PL 2338/2023 não é apenas uma norma técnica. É um divisor de águas para a governança corporativa no Brasil. Ao estabelecer a responsabilidade objetiva para sistemas de alto risco, a legislação retira das empresas o benefício da dúvida e impõe um novo padrão de diligência. Não se trata de cumprir mais um requisito regulatório: trata-se de uma mudança estrutural na arquitetura de risco de qualquer empresa que opere IA em processos decisórios com impacto patrimonial.

Para o conselho de administração e o C-Level, o risco de operar IAs de decisão — sinistros, crédito, apólices — sem autenticidade verificável dos dados de entrada é triplo e concreto:

⚠️ **Risco Financeiro**	Exposição a multas de até R$ 50 milhões ou 2% do faturamento do grupo econômico, acumulada ao custo direto de sinistros fraudulentos aprovados sobre evidências manipuladas — prejuízo que os modelos de IA atuais não conseguem interceptar sem autenticidade de input.
⚖️ **Risco Jurídico**	Em regime de inversão do ônus da prova, a incapacidade de demonstrar que a decisão da IA partiu de um dado íntegro torna a defesa judicial tecnicamente inviável. O departamento jurídico não pode construir uma linha de defesa sobre evidências cuja autenticidade não pode ser provada.
🏛️ **Risco Reputacional**	A fragilidade nos controles de auditabilidade pode ser interpretada pela ANPD como omissão negligente — com impacto direto na confiança de investidores, parceiros de resseguro e clientes no ecossistema digital da companhia. Em um setor onde a confiança é o ativo central, um auto regulatório é um evento de comunicação corporativa.

C2PA com ICP-Brasil como Seguro de Governança

A adoção do padrão C2PA, reforçado pela infraestrutura ICP-Brasil, fecha a "última milha" da auditabilidade que o mercado ainda ignora: a camada da evidência. A auditabilidade do modelo de IA — explicabilidade, logs de decisão, análise de viés — está recebendo atenção crescente. A auditabilidade de onde os dados vieram, ainda não. Essa é a lacuna que compromete toda a cadeia.

Implementar o Provvi hoje não é uma escolha técnica de TI. É uma decisão estratégica de compliance proativo — a garantia de que a inteligência artificial da sua empresa está construída sobre uma base de fatos verificáveis, e não sobre vulnerabilidades invisíveis que só se tornam visíveis no momento de uma auditoria regulatória ou de um processo judicial.

Próximo Passo para o Conselho

A Provvi está preparada para realizar uma sessão de avaliação de impacto algorítmico e demonstração de auditabilidade focada na sua esteira de decisão atual.

Objetivo: Identificar lacunas de autenticidade no pipeline de IA existente e apresentar o roteiro de integração do SDK sem impacto nos sistemas atuais.

Agende uma reunião executiva: provvi.com.br · contato@provvi.com.br

Sobre o Provvi

O Provvi é um SDK Android e iOS de captura autenticada de imagens com validade jurídica via C2PA e ICP-Brasil. Desenvolvido para setores que precisam de evidência digital defensável — seguros, vistorias, registros regulatórios e educação —, o Provvi implementa o padrão internacional de proveniência de conteúdo com assinatura via infraestrutura de chaves públicas brasileira.

Site: provvi.com.br · Contato: contato@provvi.com.br

Sobre a Autoria

Este whitepaper foi desenvolvido pelo Comitê de Tecnologia da Provvi, composto por especialistas em criptografia, infraestrutura de chaves públicas (ICP-Brasil) e direito digital. A Provvi dedica-se ao desenvolvimento de padrões de proveniência de conteúdo para setores de alta criticidade, assegurando que a inovação tecnológica em IA e automação esteja sempre amparada por evidências digitais defensáveis e conformidade regulatória plena.

Produzido por: Comitê de Tecnologia e Compliance da Provvi · Data de Publicação: Março de 2026

Referências Regulatórias

- PL 2338/2023 — Senado Federal. Dispõe sobre o uso da Inteligência Artificial. Aprovado pelo Senado em 10/12/2024; em análise na Câmara dos Deputados.
- AI Act — Regulamento (UE) 2024/1689, Parlamento Europeu e Conselho da UE, agosto de 2024.
- ICP-Brasil — Medida Provisória 2.200-2/2001, Infraestrutura de Chaves Públicas Brasileira.
- LGPD — Lei 13.709/2018, Lei Geral de Proteção de Dados Pessoais.
- C2PA Specification v2.1 — Coalition for Content Authenticity and Provenance (contentauthenticity.org).
- ANPD — Análise preliminar do PL 2338/2023. Autoridade Nacional de Proteção de Dados, 2023.
- CNseg — 22º Ciclo SQF: Seminário de Quantificação da Fraude em Seguros, 2024.