Shadow IA: O Uso Não Autorizado de Inteligência Artificial nas Corporações e seus Riscos

Autores: Marcelo Branquinho e Plataforma Safer

03/10/2025

Resumo: O uso massivo de ferramentas de inteligência artificial generativa trouxe eficiência e inovação, mas também fez emergir o fenômeno da **Shadow IA**, caracterizado pelo uso não autorizado de aplicações de IA por funcionários fora da governança corporativa. Esse comportamento, muitas vezes motivado por boas intenções de produtividade, pode resultar em **vazamento de dados sensíveis, exposição de propriedade intelectual, violação de legislações como LGPD e GDPR, e fortalecimento de ataques de engenharia social**. Casos reais, como os vazamentos na Samsung e as proibições em bancos internacionais e nacionais, evidenciam a gravidade do problema. O artigo analisa **vetores de risco, impactos operacionais, reputacionais e legais**, além dos desafios para diagnosticar esse uso invisível. Por fim, propõe **estratégias de prevenção e contenção**, incluindo políticas claras de IA, alternativas seguras, monitoramento técnico e programas de conscientização, reforçando a necessidade de uma governança robusta. A mensagem central é que a IA deve ser incorporada de forma **responsável, transparente e sob controle corporativo**, evitando que a Shadow IA se torne uma ameaça à segurança da informação e à conformidade.

Palavras-chave: Shadow IA, Shadow IT, Governança de IA, Segurança da informação, Vazamento de dados, LGPD, GDPR, Engenharia social, Compliance corporativo, Riscos cibernéticos, Inteligência artificial generativa.

1. Introdução

A adoção explosiva de ferramentas de inteligência artificial (IA) generativa nos últimos anos trouxe ganhos de produtividade, mas também um novo conjunto de riscos para a segurança corporativa. **Shadow IA** refere-se ao uso de aplicações de IA por funcionários sem aprovação ou conhecimento do departamento de TI – um análogo específico da IA para o fenômeno de *Shadow IT*. Com plataformas como ChatGPT tornando-se amplamente acessíveis, colaboradores muitas vezes recorrem a essas ferramentas por conta própria para agilizar tarefas, desconhecendo que podem expor inadvertidamente a organização a sérias ameaças de segurança da informação, conformidade e reputação.

Indicadores recentes confirmam a dimensão crescente do problema. Entre 2023 e 2024, a porcentagem de funcionários corporativos utilizando aplicações de IA generativa saltou de 74% para 96%, e **mais de um terço (38%) dos empregados admitem já ter compartilhado informações de trabalho confidenciais com ferramentas de IA sem autorização da empresa**. Além disso, um relatório sobre segurança de dados de 2025 revelou que 98% dos funcionários usam aplicativos não aprovados em cenários de Shadow IT/AI, evidenciando quão disseminado é o uso "por baixo dos panos" desses recursos. Esse quadro acende um alerta: o emprego não supervisionado de IA – a Shadow IA – pode **resultar em vazamento de dados sensíveis, violação de políticas e obrigações legais, e abertura de vulnerabilidades invisíveis no ambiente corporativo**. Em outras palavras, a IA sem controle adequado pode rapidamente deixar de ser aliada para se tornar uma ameaça à segurança da informação.

Neste artigo técnico, exploramos em detalhes o fenômeno da Shadow IA e por que ele se tornou um risco crescente. Discutiremos casos reais documentados – no Brasil e no exterior – de uso indevido de IA por funcionários ou equipes não autorizadas, analisando os vetores de risco envolvidos (desde vazamentos de dados e exposição de segredos corporativos até ataques de engenharia social e violações de compliance). Em seguida, examinaremos os impactos operacionais,

reputacionais e legais que podem advir desse cenário, incluindo possíveis infrações à LGPD, GDPR e às diretrizes de segurança corporativa. Abordaremos como as empresas podem diagnosticar a presença de Shadow IA em seus ambientes e os desafios para rastrear esse uso "invisível". Por fim, apresentaremos estratégias de prevenção e contenção – englobando políticas, processos e ferramentas – para mitigar os riscos, aliando governança de dados, segurança da informação e conscientização dos colaboradores. A mensagem central é que a IA pode e deve ser integrada de forma responsável e supervisionada nos negócios; caso contrário, a Shadow IA pode se tornar uma perigosa "porta dos fundos" para incidentes de segurança e compliance.

2. Exemplos Reais e Casos Documentados

Para compreender a gravidade da Shadow IA, é útil observar incidentes já registrados em organizações que ilustram os riscos envolvidos. Um dos casos mais emblemáticos ocorreu na **Samsung Electronics, em 2023**. Em questão de semanas, a empresa identificou pelo menos **três vazamentos de informações sensíveis causados por funcionários** que utilizavam o ChatGPT de forma inadvertida. Em um dos episódios, um engenheiro copiou código-fonte confidencial de uma base de dados de semicondutores e colou no chatbot na esperança de obter correção de erros; em outro, um funcionário inseriu código proprietário buscando otimização; o terceiro incidente envolveu um colaborador solicitando ao ChatGPT a ata de uma reunião interna, fornecendo detalhes da reunião nos prompts. Esses casos demonstraram exatamente o que especialistas vinham alertando: dados inseridos em ferramentas de IA públicas acabam fora do perímetro de controle da empresa e potencialmente integrados ao modelo, podendo ser recuperados por terceiros mal-intencionados no futuro. Diante desses incidentes, a Samsung tomou medidas emergenciais para bloquear o uso interno de IA generativa. Menos de um mês após liberar o acesso ao ChatGPT, a companhia voltou atrás e **baniu o uso de chatbots de IA em seus dispositivos e redes**

corporativas – implementando também limites estritos no tamanho das consultas e ameaçando medidas disciplinares para vazamentos futuros. Em suma, a experiência da Samsung expôs concretamente os riscos da Shadow IA e forçou uma resposta rápida de contenção.

O caso Samsung não é isolado. **Diversas grandes empresas internacionais adotaram políticas restritivas em 2023** reagindo a riscos semelhantes. Por exemplo, **a Apple proibiu seus funcionários de usar o ChatGPT** em atividades de trabalho após suspeitas de que a IA poderia reter informações proprietárias. Outros gigantes como **Amazon, JPMorgan Chase, Bank of America, Goldman Sachs e Citigroup impuseram vedações ou limites ao uso de IA generativa internamente**, temendo vazamentos de segredos comerciais e dados de clientes. Em certos casos, o bloqueio foi total; em outros, proibiu-se ao menos o uso de versões gratuitas ou não corporativas das ferramentas. A preocupação central, segundo reportado por fontes internas desses bancos, era de que empregados pudessem inserir dados confidenciais de clientes ou códigos-fonte em sistemas externos de IA, expondo tais informações a servidores fora do controle da companhia. De fato, a própria OpenAI, desenvolvedora do ChatGPT, **confirmou em 2023 uma falha que expôs históricos de conversas e dados de pagamento de usuários** devido a um bug – o que reforçou o temor de que qualquer dado fornecido ao chatbot poderia vazar em caso de brechas de segurança. Esse incidente levou até mesmo países e governos a agirem: a **Itália, por exemplo, impôs um bloqueio temporário ao ChatGPT em março de 2023** por violações de privacidade, tornando-se o primeiro país ocidental a barrar a ferramenta enquanto a OpenAI não atendesse aos requisitos do regulador nacional de dados. Em outras jurisdições, agências de proteção de dados emitiram orientações severas quanto ao uso de IA generativa, alertando empresas sobre as implicações legais de tratar dados pessoais nessas plataformas.

No **Brasil**, embora não tenham vindo a público vazamentos específicos atribuídos à Shadow IA em 2023, o movimento preventivo também aconteceu. Um levantamento feito pelo portal Tilt (UOL) junto aos cinco maiores bancos brasileiros revelou que **ao menos três deles**

já restringiam o uso do ChatGPT entre seus funcionários em meados de 2023. Em um dos bancos, permitia-se apenas a versão corporativa paga (com políticas de privacidade mais claras), enquanto outro optou pelo bloqueio completo do chatbot em redes internas. A Federação Brasileira de Bancos (Febraban) esclareceu que não há norma setorial específica no país, cabendo a cada instituição definir sua política, mas o fato é que grandes bancos nacionais se adiantaram em proibir ou limitar o uso de IA generativa, ecoando as práticas das matrizes estrangeiras. A justificativa é a mesma observada globalmente: **temor de que informações sensíveis de clientes ou dados financeiros privilegiados acabem sendo transferidos indevidamente para servidores externos** ao serem inseridos em prompts de IA. Em setores altamente regulados, como o financeiro, esse risco é considerado inaceitável, dado que viola diretrizes básicas de sigilo bancário e proteção de dados. Portanto, ainda que nem todas as empresas brasileiras tenham formalizado políticas sobre ferramentas de IA, já em 2023 algumas das mais críticas (bancos, especialmente) adotaram postura conservadora, preferindo **vetar a IA não homologada antes que ocorra um incidente de grandes proporções**.

Outros casos ilustram a ubiquidade da Shadow IA e seus perigos latentes. Uma pesquisa da empresa de segurança Cyberhaven em 2023 flagrou múltiplos exemplos de uso indevido de IA generativa por funcionários de clientes corporativos. Em um dos episódios, **um executivo sênior inseriu no ChatGPT todo o documento estratégico anual de sua empresa** para que a IA gerasse uma apresentação em slides – potencialmente expondo planos de negócios confidenciais a um sistema externo. Em outra ocorrência, **um médico forneceu a um chatbot dados sensíveis de um paciente (nome e diagnóstico)** para obter ajuda na redação de uma carta à seguradora, violando protocolos de privacidade de saúde. Esses exemplos revelam que a Shadow IA perpassa diferentes indústrias e níveis hierárquicos: do setor de tecnologia à área médica, de funcionários juniores a executivos, muitos veem nas ferramentas de IA um “atalho” conveniente e as utilizam sem aval, abrindo brechas preocupantes. Em todos os casos acima, os riscos associados – vazamento de propriedade intelectual, quebras de

sigilo profissional e exposição de dados pessoais – materializaram-se ou estiveram muito próximos de se materializar.

Em síntese, os casos reais documentados mostram que a Shadow IA deixou de ser um cenário hipotético para se tornar uma **realidade concreta nas organizações**. Grandes corporações já experimentaram na prática os danos potenciais (como a Samsung) ou identificaram riscos suficientemente alarmantes para justificar **políticas de tolerância zero** (como bancos e empresas de tecnologia). Tais eventos ressaltam a importância de reconhecer o fenômeno e endereçá-lo proativamente – antes que um incidente de Shadow IA cause prejuízos financeiros, legais ou reputacionais difíceis de reverter.

3. Vetores de Risco e Ameaças Envolvidas

A Shadow IA concentra uma combinação de vetores de risco que a torna especialmente insidiosa. Diferentemente de ameaças tradicionais, aqui os **funcionários agem sem má-fé, porém suas ações podem provocar consequências típicas de um ataque interno ou de uma falha de segurança**. Podemos destacar quatro categorias principais de ameaças relacionadas ao uso não autorizado de IA generativa: **vazamento de dados confidenciais**, **exposição de segredos corporativos/propriedade intelectual**, **facilitação de ataques de engenharia social** e **violações de compliance e privacidade**. A seguir, detalhamos cada um desses vetores.

- Vazamentos de dados e exposição de segredos corporativos: Este é o risco mais imediato. Quando um colaborador insere informações corporativas em uma ferramenta de IA pública, ele pode **estar transmitindo dados sensíveis para servidores fora do perímetro seguro da organização**. Por exemplo, imagine-se um representante de vendas que copia e cola um contrato de um cliente em um chatbot de IA para gerar um resumo dos pontos principais: sem perceber, ele possivelmente expôs *estratégias de preços confidenciais, dados de clientes e cláusulas proprietárias* a um serviço externo, fora do controle da empresa. Uma vez enviados, esses dados podem acabar armazenados ou incorporados ao modelo de IA, tornando-se parte de

sua base de treinamento. **Informações sigilosas podem, assim, "vazar" de forma inadvertida**, seja por acesso indevido de terceiros aos servidores da IA, seja por serem **reaproveitadas nas respostas do modelo a outros usuários futuramente**. Vale lembrar que a OpenAI *alerta explicitamente* que não é possível deletar prompts individuais do histórico e orienta usuários a não inserirem dados sensíveis nas conversas. Ou seja, dados internos uma vez fornecidos a um LLM (Large Language Model) público podem permanecer acessíveis de alguma forma. Esse vetor de risco inclui não apenas informações de negócio, mas também *código-fonte proprietario*: como vimos, desenvolvedores buscando ajuda para corrigir ou otimizar código podem acabar entregando trechos valiosos de software a um repositório externo (o modelo de IA), correndo o risco de violar segredos industriais. Em resumo, **a Shadow IA pode funcionar como um canal de exfiltração de dados**, ainda que não intencional – um problema gravíssimo em termos de segurança da informação.

- Engenharia social e fraudes facilitadas pela IA: Outro vetor de ameaça é a potencial exploração maliciosa dessas ferramentas, seja por funcionários despreparados ou por agentes externos, resultando em ataques de engenharia social mais eficazes. Por um lado, um colaborador desavisado pode ser induzido, por meio de respostas enganosas de um chatbot comprometido, a revelar mais informações do que deveria (*prompt injection attacks*) ou a executar ações inseguras. Por outro lado – e mais criticamente – **cibercriminosos podem usar IA generativa para potencializar ataques de phishing e fraude direcionados à empresa**, explorando dados vazados pela Shadow IA. Informações internas obtidas indevidamente (por exemplo, detalhes de um contrato, nomes de projetos ou estilo de comunicação interna) podem ser usadas para compor mensagens fraudulentas extremamente convincentes. Ferramentas de IA permitem criar textos impecáveis, personalizados e em qualquer idioma, eliminando sinais típicos de golpes amadores. Estudos mostram que **phishings elaborados com auxílio de IA tiveram uma taxa de sucesso 92% maior em enganar usuários do que ataques genéricos tradicionais**. Além disso, a IA tornou triviais os ataques de *spear phishing*

multilíngues – hoje um golpista russo ou brasileiro pode gerar e-mails quase perfeitos em inglês, francês ou mandarim, ampliando seu alcance global. Essa sofisticação aumentou drasticamente o volume de ataques: um relatório de 2024 apontou um salto de **1.265% no número de ataques de phishing após a popularização do ChatGPT**, indicando que os criminosos estão produzindo muito mais conteúdo malicioso com auxílio de IA. Consequentemente, esquemas como *Business Email Compromise (BEC)* também dispararam – +1760% em volume entre 2022 e 2024 – causando prejuízos bilionários e desafiando as defesas tradicionais. Em suma, a Shadow IA pode tanto servir de **vetor inicial (vazando dados que serão usados em golpes)** quanto de **ferramenta nas mãos de atacantes (gerando mensagens enganosas difíceis de detectar)**.

Um exemplo extremo desse tipo de ameaça são as fraudes com **deepfakes** de voz e vídeo, facilitadas por IA. Com poucos minutos de áudio original, hoje é possível clonar a voz de uma pessoa com alta fidelidade. Criminosos já exploram isso para se passar por executivos ou parceiros de negócio, enganando funcionários. Em **2019, o CEO de uma empresa de energia no Reino Unido foi induzido a transferir €220 mil** a um fornecedor falso após receber uma ligação de alguém imitando perfeitamente a voz do diretor da matriz (com sotaque e entonação realistas) – tratava-se de um áudio sintético gerado por IA. Desde então, casos similares só se multiplicaram. Em 2023, um funcionário de uma empresa em Hong Kong autorizou uma transferência de **US$ 25 milhões** após participar de uma videoconferência em que viu e ouviu o que acreditava serem seu diretor financeiro e outros colegas aprovando a operação – na verdade, eram *avatares gerados por IA* dos executivos, meticulosamente criados por golpistas. Estes casos demonstram o poder da IA em elevar a engenharia social a um patamar inédito: *não se pode mais confiar cegamente em voz ou imagem como prova de identidade*. Assim, a Shadow IA amplifica o risco de ataques de impostores contra a própria empresa, seja fornecendo munição (dados internos vazados) para golpistas, seja tornando colaboradores alvo de fraudes sofisticadas que exploram IA.

- Ameaças à privacidade e compliance (LGPD, GDPR etc.): O uso não autorizado de IA também pode resultar em **violações diretas de leis de proteção de dados e normas regulatórias**. Quando um funcionário alimenta um sistema de IA externo com dados pessoais de clientes, por exemplo, ele pode estar infringindo legislações como a **LGPD (Lei Geral de Proteção de Dados)** no Brasil e o **GDPR (General Data Protection Regulation)** na União Europeia. Essas leis exigem bases legais para tratamento de dados pessoais e restringem a transferência de dados para terceiros ou para fora do país sem salvaguardas adequadas. Inserir informações pessoais (sejam dados financeiros de um cliente, sejam dados sensíveis como estado de saúde) em um serviço de IA online sem aval jurídico pode ser interpretado como *transferência internacional de dados não autorizada* ou mesmo divulgação indevida a terceiros. **Especialistas alertam que anexar dados pessoais ou documentos confidenciais em um chatbot pode configurar violação da LGPD** – afinal, a empresa está repassando esses dados a uma ferramenta cuja política de uso pode permitir reaproveitamento das informações. No caso citado pela CNN Brasil, enviar resultados de exames médicos a um chatbot violaria diretamente a LGPD e poderia acarretar sanções. Ainda que não haja intenção maliciosa, o simples ato pode constituir descumprimento de requisitos legais de privacidade. Some-se a isso o risco de tais dados vazarem (por falha técnica ou incidente de segurança da plataforma de IA), expondo indivíduos e a empresa a consequências legais e financeiras. Em setores regulados, *qualquer* compartilhamento de informação regulatória através de canais não homologados também pode ferir normas específicas (por exemplo, dados de identificação de clientes financeiros – KYC – são protegidos por sigilo bancário). Ou seja, a Shadow IA traz consigo o **vetor de compliance**: uso indevido de IA pode levar a **descumprimento de obrigações legais e regulatórias**, sujeitando a organização a multas, processos e outras penalidades.

Resumidamente, a Shadow IA encapsula riscos de **vazamento involuntário de informações**, **fortalecimento de ataques externos de engenharia social** e **transgressão de leis e normas internas de segurança**. É uma ameaça multifacetada: ao mesmo tempo

tecnológica (por envolver plataformas de IA externas) e humana/procedimental (por decorrer de ações de pessoas fora de políticas estabelecidas). Esses vetores evidenciam por que **o uso não monitorado de IA generativa representa hoje uma das principais preocupações em segurança da informação**. A seguir, examinamos os impactos que tais riscos podem gerar se concretizados – desde danos às operações diárias da empresa até perdas reputacionais e multas regulatórias significativas.

4. Impactos Operacionais, Reputacionais e Legais

Os riscos da Shadow IA não são meramente teóricos – **suas consequências podem afetar profundamente diversos aspectos de uma organização**, desde a continuidade dos negócios até sua imagem pública e situação legal. Nesta seção, analisamos os potenciais impactos operacionais, reputacionais e legais decorrentes do uso não autorizado de IA nas corporações, com destaque para cenários envolvendo violações da LGPD, GDPR e diretrizes internas de segurança.

- Impactos operacionais: No curto prazo, um incidente de Shadow IA pode desencadear transtornos operacionais significativos. Vazamentos de informações críticas costumam exigir respostas emergenciais: investigações internas, esforços de contenção e correção, notificações a partes afetadas e eventualmente ao regulador (conforme a legislação). Esses processos desviam equipes de suas funções normais, gerando **custos adicionais e perda de produtividade**. Imagine, por exemplo, que códigos-fonte confidenciais de um produto foram expostos em um chatbot – a empresa talvez precise suspender temporariamente lançamentos ou modificar sistemas apressadamente para mitigar o impacto, incorrendo em atrasos e gastos não planejados. Mesmo quando não há vazamento confirmado, **decisões ou ações baseadas em saída de IA incorreta podem prejudicar operações**: se um funcionário confiar cegamente em uma resposta errônea de IA ao elaborar um relatório ou configurar um sistema, o resultado pode ser retrabalho, falhas de processo ou erros em produtos entregues. Modelos de IA não supervisionados

podem introduzir *viéses ou informações fictícias (alucinações)* em análises corporativas, levando a **escolhas estratégicas ruins que afetem o desempenho do negócio**. Além disso, uma vez que a Shadow IA opera fora dos controles normais, ela escapa de mecanismos de backup, auditoria e registro – logo, caso algo dê errado (por exemplo, um conteúdo gerado inadequado seja publicado ou um dado seja perdido), **não há trilha clara para entender o ocorrido rapidamente**, dificultando a resposta e retomada das operações. Em suma, do ponto de vista operacional, a Shadow IA pode gerar desde incidentes de segurança (que interrompem atividades) até decisões subótimas (que impactam resultados), minando a eficiência e a confiabilidade dos processos internos.

- Impactos reputacionais: A confiança de clientes, parceiros e do mercado é um ativo intangível vital para as empresas – e pode ser gravemente abalada por incidentes envolvendo Shadow IA. Vazamentos de dados sigilosos, em especial, costumam repercutir negativamente na opinião pública. Se ficar conhecido que a empresa expôs informações de clientes ou propriedade intelectual ao usar inadequadamente uma ferramenta de IA, **haverá danos à credibilidade e à imagem de profissionalismo da organização**. Caso stakeholders-chave (clientes, acionistas, órgãos de imprensa) descubram que decisões ou conteúdos foram produzidos por IA sem controle de qualidade, pode-se instalar a percepção de falta de rigor e governança. Por exemplo, a **Sports Illustrated sofreu forte reação negativa em 2023 ao ser revelado que publicara artigos escritos por IA como se fossem de autores humanos**, levantando questionamentos éticos e prejudicando sua reputação de integridade jornalística. Da mesma forma, o **Uber Eats enfrentou críticas ao usar imagens de refeições geradas por IA** sem informar claramente aos consumidores, o que gerou desconfiança quanto à autenticidade da plataforma. No contexto corporativo, imagine um cliente corporativo descobrindo que sua consultoria usou ChatGPT para gerar a maior parte de um relatório estratégico – a confiança no valor daquele trabalho despencaria. Concluindo, a Shadow IA pode comprometer seriamente a percepção de confiabilidade da empresa, tanto

externamente quanto junto a seus próprios gestores, ao evidenciar falta de controle sobre seus processos e informações. Reconstruir reputação perdida é uma tarefa árdua; assim, evitar esses lapsos de governança é crucial.

- Impactos legais e de compliance: Talvez a dimensão mais tangível (e potencialmente onerosa) dos impactos da Shadow IA resida nas implicações legais. Como discutido, o uso não autorizado de IA pode levar a **violação de leis de proteção de dados**, acordos contratuais e regulamentações setoriais. No caso da **LGPD**, uma infração séria – como compartilhar dados pessoais sensíveis de clientes com um serviço de IA sem base legal – pode sujeitar a empresa a multas de até R$ 50 milhões por infração, além de sanções como publicização do incidente e bloqueio dos dados envolvidos. Já no âmbito do **GDPR europeu**, as penalidades são notoriamente severas: violações graves (por exemplo, transferência ilícita de dados para fora do Espaço Econômico Europeu) podem acarretar multas de **até €20 milhões, ou 4% do faturamento global anual da companhia, prevalecendo o maior valor**. Esses montantes ilustram o impacto financeiro direto que um deslize de Shadow IA pode ter. E não se trata só de multas: a empresa pode enfrentar processos judiciais movidos por clientes ou parceiros lesados pelo vazamento de informações (*class actions*, no caso de dados pessoais expostos, são uma possibilidade em países onde há esse instrumento). Há também o risco de **penalidades contratuais** – se informações cobertas por acordos de confidencialidade (NDAs) vazam via IA, os clientes ou fornecedores afetados podem invocar cláusulas de penalidade ou rescindir contratos por quebra de sigilo. Ademais, dependendo do setor, reguladores específicos podem impor sanções administrativas adicionais. Por exemplo, instituições financeiras no Brasil sob regulação do Banco Central poderiam sofrer restrições ou intervenções se falharem em proteger dados sensíveis. Importante notar que a responsabilidade não recai apenas sobre a empresa: **o funcionário que compartilha dados protegidos via Shadow IA pode incorrer em infração disciplinar e até crime**, como o de divulgação de segredo profissional (art. 154 do Código Penal) caso fique configurado que revelou, sem justa causa,

segredos de que tinha ciência em razão do cargo. Em última instância, a empresa ainda pode ser cobrada por *omissão de controle*, ou seja, por não ter implementado salvaguardas para evitar o uso indevido de IA. Em suma, o **impacto legal** da Shadow IA abrange **multas elevadas, processos onerosos, responsabilização criminal e regulatória, além de violações contratuais**, compondo um leque de consequências potencialmente desastrosas para a organização.

Diante de todos esses possíveis danos – interrupção de operações, perdas financeiras, abalo da confiança do mercado e penalidades legais – fica claro que a Shadow IA representa um **risco multidimensional**. A ocorrência de um único incidente grave pode resultar em perdas que se propagam dessas três esferas (operacional, reputacional, legal) de forma interligada. Por exemplo: um vazamento de dados (incidente de segurança) leva à multa e processo (legal) e à exposição negativa na mídia (reputacional), afetando a base de clientes e exigindo esforços internos intensos (operacional) – tudo originado de um ato isolado de uso indevido de IA por um colaborador. Prevenir a Shadow IA, portanto, não é apenas evitar um "probleminha de TI", mas sim **proteger a continuidade do negócio, a imagem pública e a conformidade legal da empresa**.

5. Diagnóstico Corporativo: Identificando a Shadow IA

Detectar a presença da Shadow IA no ambiente corporativo é, por si só, um grande desafio. Como o nome sugere, trata-se de um uso "sombrio", **invisível aos mecanismos tradicionais de controle de TI** em muitos casos. Diferentemente de um software não autorizado instalado em um computador (que poderia ser descoberto por inventários de ativos ou scanners de rede), uma interação com um chatbot na web muitas vezes não deixa rastros claros nos sistemas internos. Ainda assim, é fundamental que as empresas busquem **mecanismos de diagnóstico** para trazer essa atividade à luz. Aqui abordamos formas de identificar Shadow IA e as dificuldades envolvidas.

Primeiramente, é importante adotar uma *postura proativa de reconhecimento*. Em outras palavras, a empresa deve **admitir que o fenômeno existe e pode estar acontecendo em suas equipes**, ainda que informalmente. Como enfatiza Saldanha, "não se controla aquilo que não se nomeia" – logo, o passo inicial é quebrar o tabu e discutir internamente a Shadow IA como um risco real. Uma vez ciente, a organização pode **mapear onde e como os funcionários estão usando IA sem aprovação**. Ferramentas práticas para isso incluem: **pesquisas e entrevistas internas**, nas quais se pergunta anonimamente aos colaboradores se usam ChatGPT/IA e para quais fins; **análise de logs de rede e firewall**, procurando acessos a domínios de IA (por exemplo, api.openai.com, bard.google.com) e volume de dados trafegados nesses acessos; e **monitoramento de tráfego web por proxies ou CASB (Cloud Access Security Broker)**, que podem identificar uso de aplicativos em nuvem não autorizados. Por exemplo, um CASB bem configurado consegue listar aplicações SaaS utilizadas na rede corporativa e destacar aquelas não sancionadas – o que incluiria provavelmente serviços de IA generativa. Outra frente de diagnóstico é avaliar **artefatos de trabalho** em busca de sinais de IA: documentos, códigos ou apresentações com trechos que aparentam ter sido gerados por IA (estilo de escrita uniforme, ausência de metadados de autoria, pequenas inconsistências típicas de respostas de chatbot). Já existem soluções de detecção de texto gerado por IA, embora não infalíveis, que poderiam apontar uso em relatórios internos.

No entanto, **há limites claros para a visibilidade que a empresa consegue ter**, o que configura o principal desafio no rastreamento da Shadow IA. Muitos funcionários utilizam essas ferramentas em dispositivos pessoais ou fora do ambiente controlado (por exemplo, no computador de casa, no celular via rede 4G, etc.), especialmente se a empresa bloqueia o acesso na rede corporativa. Assim, mesmo logs de rede interna podem **não capturar a extensão real do uso**, pois o colaborador pode simplesmente pular o firewall. Conforme observado no blog da Tripla, usar um ChatGPT não exige "driblar firewall" ou ter conhecimentos técnicos – **basta um navegador aberto (muitas vezes**

em dispositivo pessoal), colar um texto e pressionar Enter. Ou seja, a barreira para contornar os controles formais é baixíssima. Esse fator cultural e tecnológico dificulta o diagnóstico completo. Além disso, o **catálogo de ferramentas de IA cresce a cada dia**: mesmo que a empresa bloqueie as mais conhecidas, constantemente surgem novos aplicativos, muitas vezes *open source* ou acessíveis via web, que podem passar despercebidos. Bloquear ou monitorar *todas* as instâncias de IA é impraticável. Isso leva ao segundo grande desafio: a **evasão pelos usuários**. Se um colaborador acredita que a IA vai ajudá-lo no trabalho, ele tende a encontrar meios de burlar restrições. Pode usar o celular, ou redes Wi-Fi abertas, ou até disfarçar tráfego (usando VPNs). Inúmeros relatos anedóticos indicam que proibições simples geram a migração do problema, não sua eliminação.

Por fim, há o obstáculo da **consciência e cultura organizacional**. Muitos funcionários não enxergam o perigo em "apenas usar uma ferramenta online para trabalhar" e, portanto, não reportam esse uso nem veem motivo para interrompê-lo. Eles podem inclusive estar desconfiados de admitir tal prática em entrevistas internas por receio de punições. Esse clima de pouca transparência torna o diagnóstico mais complexo – *o risco cresce silenciosamente, motivado muitas vezes por boas intenções (entregar resultados melhores e mais rápido)*. Identificar Shadow IA, portanto, não é apenas um exercício técnico, mas também requer **quebrar a barreira do desconhecimento e do silêncio**.

Diante desses desafios, uma recomendação-chave é adotar uma mistura de **ferramentas tecnológicas e abordagens de gestão** para diagnosticar a Shadow IA. Tecnologicamente, investir em capacidades de monitoramento de aplicações em nuvem, DLP (Prevenção de Perda de Dados) para detectar quando informações confidenciais estão sendo copiadas para forms ou chats web, e análise de comportamento de usuários (UEBA) para notar atividades anômalas (como grandes blocos de texto sendo enviados para fora) pode gerar indícios valiosos. Já no âmbito de gestão, **criar um ambiente no qual os funcionários se sintam seguros para reportar o uso de novas ferramentas** pode trazer a Shadow IA à tona. Alguns programas de *compliance* incentivam os

colaboradores a declararem voluntariamente os apps que adotaram para melhorar o trabalho, sem punição, usando essa informação para avaliar riscos e eventualmente incorporar soluções oficialmente. O importante é **trazer a Shadow IA para a luz – torná-la visível e debatida**. Somente conhecendo onde ela está presente é que a organização poderá tomar medidas efetivas para controlá-la. Nas palavras de um relatório da Tripla, os maiores riscos são os que não se pode ver, e atualmente o uso oculto de IA se enquadra exatamente nisso.

Resumindo, diagnosticar a Shadow IA demanda vigilância contínua e mente aberta. *Continuidade no mapeamento*, uso de *ferramentas de detecção* e promoção de *visibilidade interna* são pilares desse processo. Ainda que nunca seja possível garantir detecção total (dado o caráter difuso e em rápida evolução das IAs), tais esforços pelo menos reduzem a "zona cega" e permitem que a empresa conheça a verdadeira extensão do problema para então agir sobre ele.

6. Estratégias de Prevenção e Contenção

Enfrentar a Shadow IA requer uma abordagem multidisciplinar, combinando ações de **governança, tecnologia e cultura**. Diferentemente de ameaças puramente técnicas, aqui é preciso gerenciar comportamento humano e incorporar a IA de forma segura nas políticas corporativas. Apresentamos a seguir um conjunto de estratégias recomendadas de prevenção e contenção do risco, cobrindo **políticas, processos e ferramentas** que as organizações podem adotar:

- **Definição de políticas claras de uso de IA:** A empresa deve estabelecer uma **política corporativa de IA** que deixe explícito quais ferramentas de IA os funcionários podem usar, em quais contextos e com quais restrições. Essas diretrizes devem cobrir, por exemplo, **quais tipos de dados jamais podem ser inseridos em sistemas de IA externos** (e.g. dados pessoais identificáveis, código-fonte sensível, informações confidenciais de clientes) e quais finalidades são aceitáveis ou não (por exemplo, usar IA para

ideação e rascunho pode ser permitido, mas para decisões finais ou gerar código de produção não). A política também deve **esclarecer as consequências para uso não autorizado** – sem cair em punições draconianas que desencorajem a cooperação, mas sinalizando seriedade. É importante alinhar essas regras às exigências legais e aos valores da empresa. Adicionalmente, deve-se atualizar as políticas existentes de **classificação da informação e gestão de dados** para contemplar riscos específicos da IA. Por exemplo: marcar certos documentos como "Proibido divulgar em ferramentas de IA" ou incluir nos treinamentos de classificação cenários envolvendo ChatGPT. Em síntese, estabelecer normas escritas é o primeiro passo para criar um entendimento comum e dar respaldo formal à governança de IA.

- **Estrutura de governança e envolvimento multidisciplinar:** Assim como ocorreu com a governança de TI no passado, as organizações precisam agora de **governança de IA**. Recomenda-se montar um comitê ou fórum interno envolvendo **TI, Segurança da Informação, Jurídico/Compliance, RH e líderes de negócio** para monitorar o tema. Esse grupo deve **avaliar riscos e casos de uso de IA** periodicamente, aprovar ou vetar ferramentas e fornecer orientações. Essa governança também abrange atualizar códigos de ética e manuais internos para incluir o uso responsável de IA. Altos executivos (CIOs, CISOs) devem patrocinar uma estratégia robusta de IA, incorporando iniciativas de segurança e compliance de IA no gerenciamento de riscos corporativos. Em resumo, enfrentar a Shadow IA não é tarefa só do departamento de TI – requer apoio de cima para baixo e coordenação entre diversas áreas, garantindo que a IA seja tratada com a mesma seriedade de qualquer outro processo de negócio crítico.
- **Oferecer** alternativas seguras **e aprovadas:** Uma das razões pelas quais a Shadow IA prospera é a ausência de opções oficiais. Se a empresa simplesmente proíbe toda e qualquer IA, mas a pressão por produtividade permanece, os funcionários buscarão

por conta própria soluções não sancionadas. Assim, uma estratégia eficaz é **fornecer ferramentas de IA aprovadas e seguras para uso corporativo**, de modo que os colaboradores não precisem recorrer a aplicativos "clandestinos". Isso pode incluir: contratação de versões *Enterprise* de serviços como OpenAI (que asseguram privacidade dos dados e permitem controle administrativo), implantação de instâncias on-premises ou privadas de modelos de IA para uso interno, ou adoção de plataformas de IA de código aberto dentro de um ambiente controlado. Por exemplo, equipes de desenvolvimento de software podem utilizar um **assistente de codificação hospedado internamente** (treinado apenas com dados seguros da empresa) para ajuda em tarefas de programação, sem expor código-fonte a plataformas públicas. Departamentos de marketing poderiam ter acesso a uma ferramenta de geração de texto aprovada, já configurada para não reter input. Ao **criar canais oficiais para usufruir dos benefícios da IA**, a empresa reduz o incentivo para o uso oculto. Essa abordagem deve vir acompanhada de comunicação clara: "Vocês podem usar estas ferramentas aprovadas, desde que sigam as diretrizes, e não devem usar outras sem aval". Em suma, integrar a IA de forma segura ao ambiente corporativo é melhor que bani-la completamente – até porque banimentos muitas vezes fracassam, conforme discutido.

- **Controle técnico e monitoramento do uso de IA:** Em paralelo às políticas e alternativas, é crucial implementar **ferramentas tecnológicas de contenção**. Soluções de **Data Loss Prevention (DLP)**, por exemplo, podem ser configuradas para detectar e bloquear tentativas de copiar informações confidenciais para campos de texto de aplicações web não autorizadas (atuando no nível do endpoint ou da rede). Firewalls e proxies devem atualizar seus filtros para **bloquear acesso a ferramentas de IA não aprovadas** ou, ao menos, alertar/administração quando ocorram. Um caso real citado foi o banimento do modelo *DeepSeek* em redes do governo dos EUA por questões de

segurança – seguindo essa linha, empresas podem manter uma lista dinâmica de aplicativos de IA proibidos (e atualizá-la constantemente, o que requer atenção do time de segurança). Outra camada é adotar **CASBs ou gateways de segurança na nuvem** capazes de interceptar uso de APIs de IA. Por exemplo, se um colaborador tenta usar um plug-in não autorizado que se comunica com um serviço de IA externo, o CASB poderia bloquear essa chamada. Adicionalmente, considerando a tendência do uso de dispositivos pessoais, soluções de **MDM (Mobile Device Management)** e políticas de BYOD devem ser revistas para cobrir o acesso a ferramentas de IA fora do perímetro corporativo. Algumas empresas optam por monitorar *logs* de DNS e trafego criptografado buscando padrões (embora aqui já entremos em privacidade do colaborador – é preciso equilíbrio). Em resumo, **ferramentas de segurança tradicionais precisam evoluir para ter visibilidade e controle sobre o tráfego relacionado à IA**, complementando as medidas organizacionais.

- **Programas de conscientização e treinamento:** Nenhuma política terá eficácia se as pessoas "na ponta" não a compreenderem ou valorizarem. Portanto, investir em **campanhas de conscientização sobre os riscos da Shadow IA** é indispensável. Isso envolve incluir módulos sobre uso seguro de IA nos treinamentos regulares de segurança da informação e compliance. Os funcionários devem aprender, com exemplos práticos, *o que não compartilhar com IAs* (reforçando pontos como os trazidos pela CNN Brasil: nunca inserir dados pessoais, senhas, informações de clientes, documentos internos etc. em chatbots) e quais são as possíveis consequências (mostrar casos reais como o da Samsung, por exemplo). Também é útil orientar sobre *como usar corretamente* as ferramentas aprovadas: por exemplo, sempre revisar criticamente as respostas da IA, nunca confiar cegamente sem validação humana, e reportar à TI qualquer comportamento anômalo observado no uso de IA. Empresas líderes estão incorporando simulações criativas nos

programas de conscientização – por exemplo, demonstrando um phishing "perfeito" gerado por IA para que o funcionário vivencie a dificuldade de detecção e entenda a importância de diretrizes de IA. A mensagem principal a passar é: **IA não é magia inofensiva – é uma tecnologia poderosa que requer responsabilidade no uso**. Com essa cultura reforçada, espera-se que os próprios colaboradores ajudem a identificar e evitar Shadow IA, agindo como parceiros da segurança em vez de elos fracos.

- **Supervisão e avaliação contínua de compliance:** Por fim, as estratégias de prevenção devem ser dinâmicas. Recomenda-se incorporar a questão da IA nas auditorias internas e avaliações de riscos periódicas. Por exemplo, equipes de compliance podem incluir perguntas sobre Shadow IA em entrevistas de auditoria com gestores de área: "Sua equipe utiliza alguma ferramenta de IA fora as oficiais? Para quê? Há controles?". Se a empresa for certificada em normas como ISO 27001, pode adicionar controles relativos a IA no escopo do SGSI (Sistema de Gestão de Segurança da Informação). Regulamentações futuras – a exemplo da *AI Act* em discussão na União Europeia – possivelmente trarão obrigações formais sobre uso de IA, então é prudente já estruturar compliance para atendê-las. Monitorar notícias e alertas de entidades de cibersegurança sobre novas variantes de Shadow IA (novos aplicativos populares, novos incidentes ocorridos em outras empresas) também faz parte da prevenção. **Liderança e especialistas devem manter-se atualizados** e prontos para ajustar políticas. A governança de IA sugerida anteriormente tem esse papel de revisão contínua. Em suma, prevenir Shadow IA não é um projeto com fim definido, e sim um processo contínuo de *melhoria e vigilância*, à medida que a tecnologia evolui e que o próprio negócio encontra novas formas de utilizar IA.

Adotando as medidas acima, as organizações podem **reduzir drasticamente a superfície de risco da Shadow IA**, ao mesmo tempo em que possibilitam colher os benefícios da inteligência artificial de

forma segura. Conforme enfatiza a IBM, implementar uma estratégia robusta de IA com foco em governança e segurança permite gerenciar os riscos da IA invisível enquanto se adotam seus benefícios. Ou seja, não se trata de frear a inovação, mas de **canalizá-la pelos trilhos corretos**. Em última análise, a IA deve ser uma aliada sob controle, e não um agente autônomo operando nas sombras.

7. Dez boas práticas para uso de IA dentro das corporações

Seguem abaixo as dez boas práticas para uso de IA dentro das corporações:

1. Defina políticas corporativas de IA estabelecendo de forma clara quais ferramentas podem ser utilizadas, em quais contextos e quais dados jamais devem ser compartilhados. Essa política deve incluir orientações práticas, como a proibição de inserir informações pessoais ou código sensível, além de prever consequências para o uso não autorizado.

2. Mantenha um catálogo de ferramentas aprovadas e bloqueie as não autorizadas, oferecendo alternativas seguras e corporativas de IA. Assim, colaboradores têm opções oficiais para usar sem recorrer a soluções externas. Acompanhe esse catálogo com bloqueios técnicos em proxies ou CASBs para evitar acessos não controlados.

3. Adote a classificação e minimização de dados em prompts, orientando que apenas o mínimo de informações seja enviado a uma IA e nunca dados confidenciais. Sempre que possível, use pseudonimização e rotule documentos com níveis de confidencialidade, deixando claro o que não pode ser compartilhado.

4. Implemente governança de risco e compliance alinhada a LGPD e GDPR, realizando análises de impacto de proteção de dados (DPIA) em casos de uso relevantes, registrando bases legais para tratamento e estabelecendo controles de transferência internacional de dados.

5. Utilize controles técnicos como DLP, CASB e monitoramento comportamental (UEBA) para identificar tentativas de exfiltração de informações via IA, bloquear colagem de dados sensíveis em chats externos e monitorar o uso de aplicativos não autorizados.

6. Garanta um ciclo de vida seguro dos modelos e mantenha humanos no processo decisório, com práticas de MLOps/LLMOps que incluam versionamento, avaliação contínua e rollback. Em atividades críticas, o princípio "human-in-the-loop" deve assegurar revisão humana das saídas da IA.

7. Proteja prompts, agentes e integrações incentivando higiene de prompts (sem expor segredos), mitigando ataques de prompt injection com filtros de saída e restrições de escopo, além de aplicar o princípio do menor privilégio em agentes autônomos e integrações via API.

8. Gerencie fornecedores e contratos com rigor, exigindo cláusulas que proíbam o uso dos dados para treinamento, garantindo segregação por cliente (tenant) e estabelecendo responsabilidades claras em caso de incidentes. Auditorias e testes de segurança periódicos devem fazer parte dessa gestão.

9. Invista em capacitação contínua e comunicação transparente, ensinando colaboradores a reconhecer riscos da Shadow IA e fornecendo guias práticos do que pode e não pode ser feito. Campanhas de conscientização com exemplos reais, simulações e exercícios tornam o aprendizado mais efetivo.

10. Prepare respostas a incidentes envolvendo IA com playbooks específicos para situações como vazamentos via prompt, alucinações em produção ou deepfakes. Inclua procedimentos para revogar acessos, notificar o DPO e reguladores quando necessário, e registrar lições aprendidas para atualizar políticas.

8. Conclusão

A ascensão vertiginosa da inteligência artificial generativa colocou nas mãos de profissionais comuns um poder de automação e criação antes inimaginável. Ferramentas como ChatGPT, Copilot, Bard, entre outras, hoje escrevem códigos, elaboram documentos e auxiliam decisões – muitas vezes com resultados impressionantes. Diante desse cenário, **é natural que os funcionários queiram incorporar a IA ao seu dia a dia de trabalho**. No entanto, conforme discutimos ao longo deste artigo, quando esse uso ocorre à revelia das políticas corporativas e sem supervisão, configura-se a chamada **Shadow IA**, que pode abrir uma perigosa porta dos fundos nas empresas. Diferentemente de adotar oficialmente a IA de forma planejada e segura, a Shadow IA floresce na informalidade, escapando aos controles e criando riscos ocultos consideráveis.

Recapitulando os pontos principais: vimos que a Shadow IA pode levar a vazamentos inadvertidos de informações confidenciais, exposição de propriedade intelectual e dados pessoais, fortalecimento de ataques de engenharia social contra a organização e potenciais violações legais (LGPD, GDPR, sigilo profissional, entre outros) com multas e sanções severas. Casos reais – como o dos funcionários da Samsung que vazaram código-fonte via ChatGPT – ilustram que esses perigos não são teóricos. Notamos também que detectar e controlar esse fenômeno não é trivial, dada sua natureza difusa e a facilidade de acesso às ferramentas de IA fora do alcance tradicional de TI. Ainda assim, **não agir não é uma opção**: ignorar a Shadow IA equivaleria a permitir a existência de um canal não monitorado por onde dados podem escapar e decisões podem ser tomadas sem accountability.

Por isso, enfatizamos a necessidade de uma **postura ativa de gestão e mitigação**. As empresas devem trazer a IA “para dentro de casa” de maneira responsável – definindo políticas claras, educando seus colaboradores, provendo ferramentas oficiais seguras e implementando salvaguardas técnicas. Integrar a IA com responsabilidade significa **equilibrar a inovação com a segurança**: aproveitar os ganhos de eficiência e criatividade que a IA oferece, mas

sem abrir mão do controle sobre informações e processos críticos. Esse equilíbrio só será atingido com governança efetiva. Em outras palavras, **a inteligência artificial deve operar sob os holofotes da governança corporativa, não nas sombras**.

Concluímos ressaltando que a IA em si não é inimiga – ao contrário, pode ser forte aliada competitiva. O que determina o desfecho (risco ou benefício) é **como ela é adotada**. A Shadow IA representa o cenário em que a adoção acontece de forma caótica, à margem das regras, potencializando ameaças. Já uma IA incorporada com critério pode elevar a empresa a novos patamares de desempenho, mantendo a proteção de seus ativos informacionais. Cabe às lideranças e profissionais de segurança da informação **antecipar-se a essa curva**, iluminando as áreas de sombra e provendo caminhos seguros para o uso de IA. Somente assim será possível colher as vantagens da era da inteligência artificial sem comprometer a confidencialidade, integridade e compliance que as organizações tanto prezam. Em suma, **IA, sim – mas com responsabilidade, transparência e controle**: essa é a chave para evitar que a Shadow IA se torne o calcanhar de Aquiles da segurança corporativa nos próximos anos.

Referências:

- Vijayan, J. “Samsung Engineers Feed Sensitive Data to ChatGPT, Sparking Workplace AI Warnings.” *DarkReading*, 11 abr. 2023.
- CNN Brasil. “ChatGPT: 7 informações que você não deve compartilhar com IAs.” *CNN Tecnologia*, 30 abr. 2025.
- FEEB-SC / UOL Tilt. “Grandes bancos do Brasil proíbem uso indiscriminado do ChatGPT.” 6 jun. 2023.
- Tripla (João Saldanha). “Shadow AI e os Riscos do que Não se Pode Ver.” *Blog Tripla*, 12 set. 2025.
- Varonis. “Riscos Ocultos da Shadow AI.” *Blog Varonis*, 21 jul. 2025.

- IBM (Tom Krantz et al). “O que é IA invisível?” *IBM Think Blog*, 25 out. 2024.
- AgileBlue (S. Dunlavey). “Why ChatGPT is Being Banned Internally in Large Organizations.” 16 mai. 2023.
- Reuters. “We wanted to craft a perfect phishing scam. AI bots were happy to help.” 8 nov. 2023.
- CNN Brasil. “Dados do trabalho e LGPD.” *CNN Tecnologia*, 30 abr. 2025.