Faculdade de Tecnologia de Americana "Ministro Ralph Biasi"

Curso Superior de Tecnologia em Segurança da Informação

GUILHERME BRUNO BERNARDES SANTANA

PRIVACIDADE E SEGURANÇA DE DADOS EM SISTEMAS DE RECOMENDAÇÃO COM INTELIGÊNCIA ARTIFICIAL NO E-COMMERCE

Americana, SP

2026

GUILHERME BRUNO BERNARDES SANTANA

PRIVACIDADE E SEGURANÇA DE DADOS EM SISTEMAS DE RECOMENDAÇÃO COM INTELIGÊNCIA ARTIFICIAL NO E-COMMERCE

Trabalho de Conclusão de Curso desenvolvido em cumprimento à exigência curricular do Curso Superior de Tecnologia em Segurança da Informação na área de concentração em Privacidade e Segurança de Dados em Sistemas de Informação

Orientador(a): Prof.(a) Maria Cristina Aranda

Este trabalho corresponde à versão final do Trabalho de Conclusão de Curso apresentado por Guilherme Bruno Bernardes Santana e orientado pelo(a) Prof.(a) Maria Cristina Aranda

Americana, SP

2026

FICHA CATALOGRÁFICA – Biblioteca Fatec Americana Ministro Ralph Biasi- CEETEPS Dados Internacionais de Catalogação-na-fonte

SANTANA, Guilherme Bruno Bernardes

Privacidade e segurança de dados em sistemas de recomendação com inteligência artificial no e-commerce. / Guilherme Bruno Bernardes Santana – Americana, 2026.

41f.

Estudo de caso (Curso Superior de Tecnologia em Segurança da Informação) - - Faculdade de Tecnologia de Americana Ministro Ralph Biasi – Centro Estadual de Educação Tecnológica Paula Souza

Orientadora: Profa. Dra. Maria Cristina Aranda

1. Análise de dados 2. Armazenamento de dados 3. Sistemas de informação. I. SANTANA, Guilherme Bruno Bernardes II. ARANDA, Maria Cristina III. Centro Estadual de Educação Tecnológica Paula Souza – Faculdade de Tecnologia de Americana Ministro Ralph Biasi

CDU: 681519
681.3.05
681518

Elaborada pelo autor por meio de sistema automático gerador de ficha catalográfica da Fatec de Americana Ministro Ralph Biasi.

Guilherme Bruno Bernardes Santana

Privacidade e Segurança de Dados em Sistemas de Recomendação com Inteligência Artificial no E-commerce

Trabalho de graduação apresentado como exigência parcial para obtenção do título de Tecnólogo em Segurança da Informação pelo Centro Paula Souza – Faculdade de Tecnologia de Americana – Ministro Ralph Biasi.
Área de concentração: Segurança de dados em e-commerce

Americana, 11 de junho de 2026

Banca Examinadora:

Maria Cristina Aranda (Presidente)
Doutora
FATEC Americana

Clerivaldo José Roccia (Membro)
Mestre
FATEC Americana

Paula da Fonte Sanches (Membro)
Mestre
FATEC Americana

DEDICATÓRIA

A minha mãe, pelo apoio incondicional em cada etapa desta jornada, e à minha namorada, pela paciência, pelo companheirismo e por estar ao meu lado nos momentos mais desafiadores. Esta conquista também é de vocês.

AGRADECIMENTOS

Em primeiro lugar, agradeço à minha família, pelo suporte constante, pelo incentivo e por todo o sacrifício feito para que eu pudesse chegar até aqui.

À minha namorada, pela paciência, pelo apoio emocional e por compreender cada hora dedicada a este trabalho.

À Prof.ª Maria Cristina Aranda, minha orientadora, pela dedicação, pelas contribuições precisas e pela condução cuidadosa deste trabalho. Sua orientação foi fundamental para que esta pesquisa tomasse forma.

Aos colegas de curso, pelas trocas, pelo companheirismo e por tornarem essa trajetória mais leve.

À Faculdade de Tecnologia de Americana "Ministro Ralph Biasi", pelo ensino de qualidade e pela estrutura que possibilitou o desenvolvimento desta pesquisa.

RESUMO

O presente trabalho investiga as vulnerabilidades e ameaças que afetam os sistemas de recomendação baseados em Inteligência Artificial no e-commerce, bem como os mecanismos técnicos e normativos disponíveis para mitigá-las. A personalização algorítmica tornou-se elemento estrutural do varejo digital, mas o extenso volume de dados pessoais necessário para seu funcionamento cria superfícies de exposição significativas. A pesquisa é de natureza bibliográfica e documental, com abordagem qualitativa, complementada por um estudo de caso exploratório sobre a Shopee Brasil. O referencial teórico abrange as principais arquiteturas de sistemas de recomendação, filtragem colaborativa, baseada em conteúdo e híbrida, os vetores de ataque documentados na literatura, envenenamento, inferência, roubo de modelo e ataques *adversariais e* os mecanismos de proteção disponíveis, como *Privacy by Design*, aprendizado federado, privacidade diferencial, anonimização e criptografia. Analisa-se também a incidência da Lei Geral de Proteção de Dados Pessoais (LGPD) e do Regulamento Geral de Proteção de Dados (GDPR) sobre o desenvolvimento e a operação dessas tecnologias. O estudo de caso identifica vulnerabilidades concretas na arquitetura da Shopee Brasil e examina o incidente de segurança de setembro de 2024, que resultou na exposição de dados cadastrais vinculados a chaves Pix. Os resultados evidenciam que a conformidade formal com a LGPD não equivale à robustez técnica necessária para enfrentar as ameaças identificadas, e que nenhum mecanismo de proteção isolado é suficiente. Conclui-se que o desenvolvimento responsável de sistemas de recomendação com IA exige a combinação de múltiplas camadas de proteção e que a privacidade deve ser tratada como requisito arquitetural desde a concepção do sistema, e não como obrigação de conformidade a ser atendida posteriormente.

Palavras Chave: Sistemas de recomendação; privacidade de dados; LGPD.

ABSTRACT

This study investigates the vulnerabilities and threats affecting Artificial Intelligence-based recommendation systems in e-commerce, as well as the technical and regulatory mechanisms available to mitigate them. Algorithmic personalization has become a structural element of digital retail, but the extensive volume of personal data required for its operation creates significant exposure surfaces. The research adopts a bibliographic and documentary approach, with qualitative analysis, complemented by an exploratory case study of Shopee Brazil. The theoretical framework covers the main recommendation system architectures, collaborative filtering, content-based filtering, and hybrid models, the attack vectors documented in the literature, poisoning, inference, model stealing, and adversarial attacks, and available protection mechanisms, including Privacy by Design, federated learning, differential privacy, anonymization, and cryptography. The impact of Brazil's General Data Protection Law (LGPD) and the European General Data Protection Regulation (GDPR) on the development and operation of these technologies is also examined. The case study identifies concrete vulnerabilities in Shopee Brazil's architecture and analyzes the security incident of September 2024, which resulted in the exposure of registration data linked to Pix keys. The findings show that formal compliance with the LGPD does not equate to the technical robustness required to address identified threats, and that no single protection mechanism is sufficient. The study concludes that the responsible development of AI-based recommendation systems requires the combination of multiple layers of protection, and that privacy must be treated as an architectural requirement from the system's conception rather than a compliance obligation to be addressed afterward.

Keywords: *Recommendation systems; data privacy; LGPD.*

LISTA DE ABREVIATURAS E SIGLAS

ABNT Associação Brasileira de Normas Técnicas

ANPD Autoridade Nacional de Proteção de Dados

CF Filtragem Colaborativa (Collaborative Filtering)

GDPR Regulamento Geral de Proteção de Dados (General Data Protection Regulation)

IA Inteligência Artificial

LGPD Lei Geral de Proteção de Dados Pessoais

MF Fatoração de Matrizes (Matrix Factorization)

ML Aprendizado de Máquina (Machine Learning)

OWASP Open Web Application Security Project

PbD Privacidade por Design (Privacy by Design)

PII Informações de Identificação Pessoal (Personally Identifiable Information)

RS Sistema de Recomendação (Recommender System)

TLS Transport Layer Security

SUMÁRIO

1 INTRODUÇÃO

Quem acessa hoje qualquer grande loja virtual raramente percebe, mas há um sistema funcionando continuamente nos bastidores: algoritmos que observam cada clique, cada produto visualizado, cada compra realizada, para montar uma vitrine personalizada. Essa é a função dos sistemas de recomendação baseados em Inteligência Artificial, tecnologia que deixou de ser diferencial para se tornar parte estrutural do varejo digital.

O impacto econômico é inegável. Bobadilla *et al*. (2013) já chamavam atenção para a dependência crescente das plataformas em relação a esses mecanismos, e os anos seguintes apenas confirmaram essa tendência. Empresas como Amazon, Mercado Livre e Shopee estruturam boa parte de sua estratégia comercial sobre a capacidade dos algoritmos de antecipar preferências e reter a atenção do consumidor.

Do ponto de vista técnico, esse nível de personalização tem um custo que frequentemente passa despercebido pelo usuário: a extensa coleta de dados pessoais. Para que um sistema de recomendação funcione adequadamente, ele precisa conhecer, ou inferir, gostos, rotinas, faixa de renda, localização e padrões de comportamento. Quanto mais dados são armazenados e processados, maior a superfície de exposição a riscos de segurança e violações de privacidade.

Esse cenário não escapou à atenção dos legisladores. A aprovação do Regulamento Geral de Proteção de Dados (GDPR) na União Europeia, em 2016, e da Lei Geral de Proteção de Dados (LGPD) no Brasil, em 2018, marcou uma virada no modo como o tratamento de dados pessoais é regulado. As obrigações impostas por essas normas interferem diretamente na forma como os sistemas de recomendação podem ser projetados e operados.

Do lado técnico, a literatura de segurança da informação vem documentando um conjunto específico de ameaças que incidem sobre esses sistemas: ataques de envenenamento que corrompem o treinamento dos modelos; técnicas de inferência que permitem reconstruir dados privados a partir dos parâmetros do algoritmo; ataques de roubo de modelo *(model stealing)*; ataques adversariais; e riscos de reidentificação que comprometem bases de dados aparentemente anonimizadas.

Diante desse contexto, este trabalho parte do seguinte problema de pesquisa: em que medida os mecanismos de proteção à privacidade adotados por plataformas de e-commerce brasileiras estão alinhados com as recomendações técnicas da literatura e com as exigências da LGPD? Para explorar essa questão, foi selecionada a Shopee Brasil como objeto de análise, por ser uma das maiores plataformas de varejo digital em atividade no país e por já ter registrado um incidente de segurança documentado em 2024.

O objetivo geral desta pesquisa é investigar as vulnerabilidades e ameaças que afetam os sistemas de recomendação com IA no e-commerce, bem como identificar os mecanismos, técnicos e normativos, capazes de reduzi-las, contrastando esse panorama com a realidade observada no caso Shopee.

Os objetivos específicos são: (a) descrever as principais arquiteturas de sistemas de recomendação utilizadas em plataformas de varejo digital; (b) mapear as ameaças à segurança e as vulnerabilidades à privacidade mais relevantes nesses sistemas, incluindo ataques adversariais e de roubo de modelo; (c) discutir abordagens de proteção como aprendizado federado, privacidade diferencial e anonimização; (d) analisar de que forma a LGPD e o GDPR incidem sobre o desenvolvimento e a operação dessas tecnologias; e (e) aplicar esse referencial ao caso Shopee Brasil, avaliando alinhamentos e lacunas.

Do ponto de vista metodológico, trata-se de uma pesquisa bibliográfica e documental, de abordagem qualitativa, complementada por um estudo de caso exploratório. O corpus de análise inclui artigos científicos, livros técnicos, documentos normativos nacionais e internacionais, a política de privacidade da plataforma e registros públicos do incidente de 2024. Esta pesquisa apresenta limitações metodológicas importantes. A análise desenvolvida baseou-se predominantemente em revisão bibliográfica, documentação pública e observação indireta do funcionamento da plataforma analisada. Não houve acesso à arquitetura proprietária, aos modelos computacionais internos ou às bases de dados da Shopee Brasil. Dessa forma, as conclusões possuem caráter analítico e interpretativo, fundamentadas em evidências publicamente disponíveis e na literatura especializada da área.

O trabalho divide-se em seis capítulos. O capítulo 2 trata dos conceitos e modelos de sistemas de recomendação. O capítulo 3 aborda os riscos à privacidade

e à segurança dos dados, com ênfase em ataques pouco explorados na versão anterior, como roubo de modelo e ataques adversariais. O capítulo 4 discute e compara as ferramentas de proteção disponíveis. O capítulo 5 apresenta a análise aplicada da Shopee Brasil. O capítulo 6 traz as considerações finais.

2 SISTEMAS DE RECOMENDAÇÃO EM E-COMMERCE

2.1 CONCEITO E EVOLUÇÃO HISTÓRICA

Um sistema de recomendação é, em essência, um mecanismo de filtragem de informação. Diante do volume cada vez maior de produtos, conteúdos e serviços disponíveis nas plataformas digitais, esses sistemas assumem a função de selecionar, entre milhares de opções, aquelas com maior probabilidade de interessar a cada usuário específico. Para Ricci, Rokach e Shapira (2011), a tarefa central desses sistemas é produzir sugestões relevantes, seja com base em padrões coletivos de comportamento, seja a partir das características dos itens em si.

As raízes dessa tecnologia remontam ao início dos anos 1990. O GroupLens, descrito por Resnick *et al.* (1994), é considerado um dos precursores: tratava-se de um sistema de filtragem colaborativa desenvolvido para auxiliar usuários do sistema de notícias Usenet a encontrar artigos relevantes entre os milhares publicados diariamente. Em paralelo, empresas de varejo americanas, com destaque para a Amazon, começavam a construir seus primeiros motores de recomendação baseados em dados de avaliação e histórico de compras.

A transformação mais significativa desse campo veio com a popularização das técnicas de aprendizado de máquina e, mais recentemente, do aprendizado profundo. Zhang *et al*. (2019) documentam como essas abordagens ampliaram a capacidade dos sistemas de capturar relações sutis entre usuários, produtos e contextos. Atualmente, os modelos processam não apenas avaliações e histórico de compras, mas também dados de localização, horário de acesso, sequência de cliques e tempo de permanência em cada página.

No âmbito do e-commerce brasileiro, esse processo acompanhou o crescimento acelerado do setor. O advento de plataformas como Mercado Livre, B2W e Shopee criou a demanda por sistemas capazes de operar em escala nacional, com milhões de produtos e dezenas de milhões de usuários ativos. Segundo a ABComm, o faturamento do e-commerce no Brasil superou R$ 185 bilhões em 2023, evidenciando o peso estratégico das tecnologias de personalização para o setor.

2.2 PRINCIPAIS MODELOS DE SISTEMAS DE RECOMENDAÇÃO

A literatura especializada organiza os sistemas de recomendação em torno de três abordagens principais, cada uma com lógica de funcionamento distinta e implicações próprias para a privacidade dos dados tratados.

2.2.1 Filtragem Colaborativa

A Filtragem Colaborativa (CF) parte de uma premissa intuitiva: pessoas que concordaram no passado tendem a concordar no futuro. Com base nisso, o sistema identifica usuários com perfis semelhantes e usa o comportamento de uns para fazer inferências sobre os outros, sem analisar o conteúdo dos itens. Koren, Bell e Volinsky (2009) foram centrais na consolidação dessa abordagem, especialmente ao demonstrar a eficácia da Fatoração de Matrizes para extrair padrões latentes de grandes bases de dados.

Do ponto de vista da privacidade, a CF apresenta uma vulnerabilidade estrutural: como aprende a partir de dados de múltiplos usuários de forma conjunta, é possível conduzir ataques de inferência que reconstituem informações sobre um usuário específico a partir dos parâmetros aprendidos pelo modelo, mesmo quando os dados originais não estão diretamente acessíveis (ZHANG; CORMODE; MIKLAU, 2021).

2.2.2 Filtragem Baseada em Conteúdo

Ao contrário da CF, a Filtragem Baseada em Conteúdo não depende do comportamento de outros usuários. Ela parte das características dos próprios itens — categoria, preço, descrição, atributos técnicos, e as compara com o perfil de preferências construído a partir das interações anteriores de cada usuário individualmente. Conforme Lops, De Gemmis e Semeraro (2011), o modelo essencialmente aprende o que o usuário aprecia e usa esse conhecimento para filtrar o catálogo.

No contexto de e-commerce, essa abordagem beneficia-se de técnicas de Processamento de Linguagem Natural para extrair representações semânticas de descrições de produtos. Em tese, protege melhor a privacidade no sentido colaborativo, mas apresenta o problema da *bolha de filtro*: sugestões excessivamente similares ao que o usuário já consumiu, empobrecendo a diversidade.

2.2.3 Sistemas Híbridos

Na prática comercial, as plataformas mais maduras combinam as duas abordagens anteriores. Conforme Burke (2002) sistematizou, é possível ponderar os resultados de dois sistemas paralelos, usar um para compensar fraquezas do outro, ou fundir ambos em um único modelo integrado.

O avanço seguinte veio com a incorporação de dados contextuais, horário, localização geográfica, dispositivo utilizado, e com arquiteturas de aprendizado profundo, como redes neurais recorrentes e modelos baseados em Transformers. O ganho em capacidade preditiva é real, mas os modelos tornam-se menos interpretáveis, dificultando a auditoria técnica e o cumprimento das obrigações de transparência da LGPD e do GDPR.

2.3 APLICAÇÕES NO E-COMMERCE

A presença dos sistemas de recomendação em plataformas de varejo digital vai muito além da clássica seção “quem comprou isso também comprou aquilo”. Eles atuam de forma distribuída ao longo de toda a jornada do consumidor: na página inicial, personalizando os destaques; nas páginas de produto, sugerindo itens complementares; no checkout, recomendando adições de última hora; e nos canais de comunicação, e-mails, notificações push, anúncios, exibindo ofertas calibradas ao perfil de cada usuário.

Linden, Smith e York (2003) descrevem a arquitetura *item-to-item collaborative filtering* da Amazon como um marco histórico: o sistema demonstrou ser capaz de operar em escala industrial, com milhões de produtos e usuários, sem perder relevância nas sugestões. Essa abordagem foi amplamente replicada pelo setor, incluindo plataformas como a Shopee, que combinam filtragem colaborativa, baseada em conteúdo e dados contextuais em tempo real para gerar recomendações personalizadas.

3 PRIVACIDADE E SEGURANÇA DE DADOS EM SISTEMAS DE IA

3.1 CONCEITOS FUNDAMENTAIS DE PRIVACIDADE DE DADOS

Falar em privacidade de dados no contexto digital exige reconhecer que estamos diante de um conceito pensado para um mundo analógico e que precisou ser reinterpretado diante de realidades novas. Westin (1967) formulou uma das definições mais citadas da área ao caracterizar privacidade como o direito do indivíduo de decidir por si mesmo quando, como e em que medida informações a seu respeito chegarão ao conhecimento de outros. Essa definição, feita décadas antes da internet, continua sendo referência por capturar a ideia central: controle sobre as próprias informações.

No ambiente digital, esse controle se torna muito mais difícil de exercer. O rastro gerado por qualquer interação online alimenta bancos de dados que, analisados por algoritmos de IA, podem revelar muito mais do que o usuário conscientemente compartilhou. Dados aparentemente inofensivos isoladamente adquirem caráter sensível quando combinados e analisados em conjunto, fenômeno que a literatura denomina risco de inferência por agregação de dados.

Na área de Segurança da Informação, a proteção desses dados articula-se com os três pilares da tríade CIA: Confidencialidade (acesso restrito a quem é de direito), Integridade (garantia de que os dados não foram adulterados sem autorização) e Disponibilidade (acesso quando necessário). Sistemas de recomendação com IA apresentam desafios em todas essas dimensões.

É importante destacar que, em sistemas de recomendação com IA, a tríade CIA apresenta tensões específicas que a tornam mais complexa do que em sistemas de informação convencionais. A Confidencialidade é desafiada pelo fato de que o próprio funcionamento do algoritmo, suas saídas públicas, pode ser explorado para inferir dados privados (ataques de inferência). A Integridade é ameaçada pelos ataques de envenenamento, que corrompem os dados de treinamento sem necessariamente violar controles de acesso. A Disponibilidade pode ser comprometida por ataques adversariais que degradam a qualidade das recomendações, afetando diretamente a experiência do usuário e a confiança na plataforma. Essas especificidades justificam

abordagens de segurança adaptadas ao contexto de IA, que vão além dos controles tradicionais de segurança da informação.

A LGPD, Lei nº 13.709/2018 (BRASIL, 2020) diferencia dados pessoais, 'informação relacionada a pessoa natural identificada ou identificável', de dados pessoais sensíveis, que incluem informações sobre saúde, origem racial, convicções religiosas e vida sexual. Sistemas de recomendação podem transformar dados classificados como anônimos em dados pessoais ao agregar e cruzar informações de múltiplas fontes.

3.2 AMEAÇAS E VULNERABILIDADES NOS SISTEMAS DE RECOMENDAÇÃO

Os sistemas de recomendação com IA combinam duas características que formam uma combinação delicada do ponto de vista da segurança: processam dados pessoais em larga escala e dependem de modelos cujo funcionamento interno pode ser explorado por adversários. As ameaças identificadas na literatura dividem-se nas categorias descritas a seguir.

3.2.1 Ataques de Envenenamento (Poisoning Attacks)

O processo de treinamento de um modelo de recomendação depende de dados históricos. Se esses dados forem corrompidos, o modelo aprende padrões distorcidos, e as recomendações geradas refletem essa distorção. Os ataques de envenenamento exploram exatamente isso: a injeção de registros fabricados no conjunto de treinamento com o objetivo de manipular as saídas do sistema a favor de interesses específicos.

Lam e Riedl (2004) foram os primeiros a sistematizar esse tipo de ataque no contexto de filtragem colaborativa, descrevendo-o como *profile injection attack*. Os autores demonstraram que mesmo um volume pequeno de perfis falsos, se estrategicamente construídos, é suficiente para alterar significativamente o comportamento do algoritmo. As motivações variam: elevação artificial de produtos nas recomendações, sabotagem de concorrentes ou desestabilização da confiança dos usuários.

3.2.2 Ataques de Inferência (Inference Attacks)

Uma propriedade indesejada que modelos de aprendizado de máquina tendem a exibir é a memorização de exemplos de treinamento. Adversários que exploram

essa característica conseguem extrair informações privadas do modelo sem jamais acessar o banco de dados original.

Shokri *et al*. (2017) formalizaram os chamados ataques de inferência de associação (*membership inference attacks*), demonstrando ser possível determinar, com alta confiança, se um dado específico integrou o conjunto de treinamento. No contexto de e-commerce, um adversário poderia descobrir que um usuário adquiriu determinado produto — um medicamento, um livro de tema delicado, um item de uso íntimo, sem jamais acessar diretamente os registros da plataforma. Os ataques de inversão de modelo permitem, adicionalmente, reconstruir aproximações dos dados originais de treinamento a partir dos parâmetros do algoritmo.

3.2.3 Ataques de Roubo de Modelo (Model Stealing Attacks)

Os ataques de roubo de modelo representam uma ameaça frequentemente subestimada em sistemas de recomendação. Nesse tipo de ataque, o adversário não precisa acessar os dados de treinamento nem os parâmetros internos do modelo: basta realizar consultas sistemáticas à API do sistema de recomendação e analisar as respostas para reconstruir um modelo funcionalmente equivalente ao original (TRAMÈR *et al*., 2016).

Do ponto de vista da segurança, as implicações são duplas. Em primeiro lugar, o modelo roubado pode ser utilizado para realizar outros ataques, como os de inferência descritos anteriormente, em ambiente controlado, sem gerar alertas nos sistemas de detecção da plataforma. Em segundo lugar, o próprio modelo representa um ativo de propriedade intelectual de alto valor; sua reprodução não autorizada configura tanto um problema de segurança quanto jurídico. Plataformas como a Shopee, que expõem sistemas de recomendação por meio de aplicativos e APIs públicas, são particularmente vulneráveis a esse vetor de ataque.

3.2.4 Ataques Adversariais (Adversarial Attacks)

Os ataques adversariais exploram a sensibilidade dos modelos de aprendizado profundo a perturbações mínimas e cuidadosamente calculadas nos dados de entrada. No contexto de sistemas de recomendação baseados em deep learning, um adversário pode manipular atributos de um produto — como a descrição textual ou as imagens associadas, de forma imperceptível para o olho humano, mas suficiente para

alterar drasticamente a posição desse item nas recomendações geradas pelo algoritmo (CHEN *et al*., 2019).

A gravidade dessa ameaça é amplificada pelo fato de que os modelos mais modernos, como aqueles baseados em Transformers, são especialmente suscetíveis a esse tipo de perturbação. Em plataformas com catálogos de milhões de produtos, é praticamente impossível auditar manualmente cada item em busca de modificações adversariais. Do ponto de vista da privacidade, ataques adversariais bem-sucedidos podem ser usados para direcionar usuários específicos a produtos fraudulentos ou coletar informações sobre o comportamento dos algoritmos de detecção de fraude.

3.2.5 Vazamento de Dados e Criação de Perfis Invasivos

Além dos vetores de ataque específicos dos modelos de IA, os sistemas de recomendação estão sujeitos ao conjunto tradicional de riscos de segurança, violações de banco de dados, falhas de autenticação, ausência de controles de acesso adequados. O agravante nesses sistemas é que uma única brecha pode expor não apenas dados cadastrais, mas perfis comportamentais detalhados de milhões de consumidores.

Há ainda o fenômeno da privacidade por agregação: dados que, isoladamente, parecem inofensivos, frequência de acesso a categorias de produtos, horário habitual de compras, termos pesquisados, revelam informações altamente sensíveis quando tratados em conjunto por algoritmos de IA. Cadwalladr e Graham-Harrison (2018) documentaram como dados de comportamento digital foram usados para inferir orientações políticas e religiosas de usuários em escala industrial, evidenciando que o perímetro real de exposição de dados em sistemas com IA é muito mais amplo do que aparenta na superfície.

3.3 LEGISLAÇÃO APLICÁVEL: LGPD E GDPR

A regulação jurídica da proteção de dados pessoais passou por uma virada importante na segunda metade da década de 2010. O GDPR europeu, aprovado em 2016 e em vigor a partir de 2018, estabeleceu um novo padrão ao fixar princípios claros para o tratamento de dados e conceder aos titulares direitos efetivos sobre suas informações. O Brasil seguiu com a LGPD, aprovada em 2018 e plenamente vigente desde 2021.

O GDPR estrutura o tratamento de dados em torno de princípios como finalidade, necessidade, transparência, segurança e responsabilização. Para os sistemas de recomendação, o princípio da minimização de dados é particularmente desafiador: exige que apenas os dados estritamente necessários para a finalidade declarada sejam coletados, o que contrasta com a lógica dominante no setor. O artigo 22 trata das decisões automatizadas e da elaboração de perfis, estabelecendo o direito do titular de não ser submetido a decisões que o afetem significativamente com base apenas em processamento automatizado.

A LGPD é construída sobre bases similares. O tratamento de dados pessoais só pode ser realizado quando amparado em uma das hipóteses legais do artigo 7º, das quais o consentimento e o legítimo interesse são as mais relevantes para o e-commerce. A Autoridade Nacional de Proteção de Dados (ANPD) pode aplicar multas de até 2% do faturamento da pessoa jurídica no Brasil, limitadas a R$ 50 milhões por infração.

Para os sistemas de recomendação, a conformidade com a LGPD implica: informar o usuário claramente sobre o uso de seus dados para personalização; disponibilizar mecanismos de revogação de consentimento e exclusão de dados; conduzir Avaliações de Impacto à Proteção de Dados (AIPD) para operações de alto risco; e garantir que os modelos possam ser auditados quanto à origem e ao uso dos dados que os alimentam.

4 MECANISMOS DE PROTEÇÃO À PRIVACIDADE EM SISTEMAS DE RECOMENDAÇÃO

4.1 PRIVACY BY DESIGN

A ideia central do *Privacy by Design* (PbD) é que a privacidade não deve ser pensada como requisito de conformidade a ser atendido após o sistema estar pronto. Ela precisa ser considerada desde a concepção da arquitetura. Cavoukian (2010) sistematizou esse princípio em sete fundamentos que abrangem desde a proatividade no tratamento de riscos até a garantia de funcionalidade plena mesmo com as proteções ativadas.

No desenvolvimento de sistemas de recomendação para e-commerce, adotar o PbD significa tomar decisões que podem não ser óbvias do ponto de vista da performance: definir com precisão quais dados são indispensáveis para cada função de recomendação e descartar o restante; estabelecer períodos de retenção e rotinas de descarte seguro; implementar controles de acesso granulares baseados no princípio do menor privilégio; e criar mecanismos que permitam ao usuário visualizar, corrigir e excluir seus dados. A abordagem está alinhada com as exigências da LGPD e é explicitamente reconhecida pelo GDPR em seu artigo 25.

4.2 APRENDIZADO FEDERADO

O aprendizado federado resolve uma limitação fundamental dos sistemas tradicionais: a necessidade de centralizar dados. No modelo proposto por McMahan *et al.* (2017), em vez de trazer os dados até o modelo, o modelo vai até os dados. Cada dispositivo de usuário treina localmente com suas próprias informações e transmite ao servidor central apenas as atualizações dos parâmetros, nunca os dados brutos.

Chai *et al*. (2020) demonstraram resultados comparáveis aos sistemas centralizados com ganhos expressivos em privacidade ao propor um framework de filtragem colaborativa federada. A ressalva importante é que o aprendizado federado não elimina todos os riscos: os gradientes transmitidos podem, em certas condições, revelar informações sobre os dados locais. Por isso, sua combinação com privacidade diferencial é especialmente recomendada.

Apesar das vantagens relacionadas à privacidade, o aprendizado federado também apresenta limitações práticas relevantes. A diferença de capacidade entre os dispositivos dos usuários pode afetar o desempenho do treinamento distribuído, principalmente em plataformas com grande volume de acessos simultâneos. Além disso, o envio constante de atualizações entre dispositivos e servidor central aumenta o custo de comunicação da infraestrutura. Em plataformas de grande porte, como a Shopee Brasil, esses fatores podem impactar tanto a eficiência operacional quanto a qualidade final das recomendações apresentadas aos usuários.

4.3 PRIVACIDADE DIFERENCIAL

A privacidade diferencial resolve um problema persistente nas abordagens de anonimização tradicionais: mesmo quando dados identificadores são removidos, os modelos treinados podem vazar informações sobre indivíduos. Dwork *et al*. (2006) propuseram introduzir ruído calibrado matematicamente no processo de aprendizado, de forma que a participação ou ausência de qualquer indivíduo no conjunto de treinamento produza efeitos indistinguíveis nas saídas do modelo.

Formalmente, um mecanismo é (ε, δ)-diferencialmente privado quando a diferença na probabilidade de qualquer saída, causada pela inclusão ou exclusão de um único registro, é controlada pelos parâmetros ε e δ. Valores pequenos de ε representam proteção mais forte, mas, geralmente, menor utilidade dos dados. Berlioz *et al*. (2015) demonstraram que é possível aplicar privacidade diferencial a sistemas baseados em fatoração de matrizes sem comprometer excessivamente a qualidade das sugestões.

4.4 ANONIMIZAÇÃO E PSEUDONIMIZAÇÃO

Anonimizar dados significa aplicar transformações irreversíveis que inviabilizem a identificação dos indivíduos. Pseudonimizar é um processo reversível: os identificadores são substituídos por códigos, mas a correspondência original é mantida em local separado. A LGPD trata essas modalidades de forma distinta: dados pseudonimizados continuam sendo dados pessoais para fins legais; dados genuinamente anonimizados saem do escopo de proteção da lei.

A eficácia da anonimização é frequentemente superestimada. Narayanan e Shmatikov (2008) cruzaram uma base supostamente anonimizada da Netflix com registros públicos do IMDb e conseguiram reidentificar usuários com alta precisão. A

lição é clara: a anonimização não pode ser tratada como solução definitiva em ambientes onde os dados coletados são abundantes e passíveis de cruzamento com outras fontes. Técnicas como k-anonimato, l-diversidade e t-closeness oferecem garantias mais robustas, mas precisam ser periodicamente reavaliadas à luz de novas técnicas de reidentificação.

Embora técnicas como k-anonimato, l-diversidade e t-closeness sejam amplamente utilizadas para reduzir riscos de identificação, elas apresentam limitações relevantes em ambientes com grande volume de dados comportamentais. Em sistemas de recomendação, diferentes informações aparentemente simples podem ser cruzadas para permitir reidentificação indireta dos usuários. Por esse motivo, a anonimização não deve ser considerada uma solução isolada para proteção da privacidade, mas sim uma camada complementar dentro de uma estratégia mais ampla de segurança e governança de dados.

4.5 CRIPTOGRAFIA E CONTROLE DE ACESSO

Para dados em trânsito, o protocolo TLS em versões atualizadas é o padrão mínimo aceitável. Em cenários de aprendizado federado, a criptografia homomórfica permite que operações matemáticas sejam executadas sobre dados ainda criptografados, eliminando a necessidade de descriptografá-los durante o processamento, embora o custo computacional ainda seja elevado.

No âmbito organizacional, o controle de acesso baseado em papéis (RBAC) e o princípio do menor privilégio são essenciais para conter o risco interno. A OWASP disponibiliza diretrizes amplamente reconhecidas para segurança de APIs e aplicações web que se aplicam diretamente aos sistemas de recomendação. A adoção de sistemas de detecção de intrusão (IDS) e plataformas de gerenciamento de eventos de segurança (SIEM) garante visibilidade sobre o ambiente, permitindo resposta ágil a incidentes.

4.6 ANÁLISE COMPARATIVA DOS MECANISMOS DE PROTEÇÃO

Os mecanismos de proteção apresentados possuem características, limitações e custos distintos. A tabela 1 a seguir sintetiza essa comparação para facilitar a tomada de decisão por parte dos arquitetos de sistemas de recomendação:

Tabela 1: Mecanismos de proteção

Mecanismo	Principal proteção	Limitação	Custo de implementação
Privacy by Design	Reduz coleta e exposição de dados desde a concepção	Exige redesenho arquitetural; pode reduzir features	Médio-Alto
Federated Learning	Dados brutos não saem do dispositivo do usuário	Gradientes podem vazar dados; comunicação pesada	Alto
Privacidade Diferencial	Garante matematicamente que indivíduos não são rastreáveis	Perda de precisão nas recomendações (trade-off ε)	Médio
Anonimização/Pseudonimização	Reduz identificabilidade nos dados armazenados	Vulnerável a ataques de reidentificação e linkage	Baixo-Médio
Criptografia + RBAC	Protege dados em trânsito e limita acesso interno	Não protege contra ataques ao modelo em si	Médio

Fonte: Próprio autor

Observa-se que nenhum mecanismo de proteção, quando aplicado isoladamente, consegue reduzir todos os riscos relacionados aos sistemas de recomendação.

A abordagem mais robusta combina privacidade diferencial e aprendizado federado para proteger os modelos, anonimização para os dados armazenados, criptografia para a infraestrutura, e Privacy by Design como orientação arquitetural. Essa combinação tende a apresentar maior alinhamento com os princípios previstos na LGPD.

4.7 ARQUITETURA DE SEGURANÇA PARA SISTEMAS DE RECOMENDAÇÃO

Os mecanismos de proteção apresentados nas seções anteriores podem ser utilizados de forma integrada em arquiteturas voltadas à segurança e privacidade de sistemas de recomendação. A proposta apresentada a seguir reúne práticas discutidas na literatura e medidas compatíveis com os riscos identificados ao longo deste trabalho.

A Camada 1, Coleta e Minimização de Dados, implementa os princípios do Privacy by Design na fase de ingestão. Apenas os dados estritamente necessários para a função de recomendação são coletados; atributos desnecessários são descartados na origem (data minimization). Identificadores diretos são imediatamente pseudonimizados antes de qualquer processamento, e políticas de retenção automática garantem o descarte seguro após o período definido. Esta camada reduz a superfície de exposição de dados brutos e é diretamente exigida pelo princípio da necessidade da LGPD.

A Camada 2, Treinamento Privado, substitui o modelo centralizado tradicional por uma arquitetura de Aprendizado Federado com Privacidade Diferencial. Os dados dos usuários permanecem nos dispositivos; apenas gradientes são transmitidos ao servidor central. Mecanismos de privacidade diferencial (ε-DP) com calibração adaptativa do parâmetro ε são aplicados sobre os gradientes antes da agregação, impedindo que a participação de indivíduos seja inferida a partir dos parâmetros do modelo. Esta abordagem mitiga diretamente os ataques de inferência de associação (membership inference attacks) e os riscos de reidentificação documentados no Capítulo 3.

A Camada 3, Proteção da Infraestrutura, cobre a segurança dos dados em repouso e em trânsito. Dados armazenados são cifrados com AES-256; comunicações entre

componentes utilizam TLS 1.3. O controle de acesso baseado em papéis (RBAC) com princípio do menor privilégio limita o acesso a dados e parâmetros de modelo apenas às identidades e processos que os necessitam. Rate limiting e monitoramento de padrões anômalos de consulta à API detectam tentativas de roubo de modelo por extração sistemática. Sistemas SIEM integram os logs de todas as camadas para resposta ágil a incidentes, mitigando o vetor de falha evidenciado no incidente da Shopee em 2024.

A Camada 4, Validação e Auditoria Contínua, garante a integridade do sistema ao longo do tempo. Pipelines automatizados de detecção adversarial analisam novas inserções no catálogo de produtos em busca de perturbações calculadas. Rotinas periódicas de auditoria verificam desvios nos padrões de recomendação que possam indicar envenenamento do modelo. Avaliações de Impacto à Proteção de Dados (AIPD) são realizadas a cada mudança arquitetural relevante, em conformidade com o artigo 38 da LGPD.

A integração dessas quatro camadas não elimina todos os riscos, o trade-off entre privacidade e utilidade das recomendações permanece um desafio real, especialmente na calibração do parâmetro ε da privacidade diferencial. Contudo, ela reduz significativamente a superfície de ataque e estabelece rastreabilidade para auditoria regulatória. A Figura 1 ilustra esquematicamente a relação entre as camadas e os vetores de ameaça que cada uma endereça.

Figura 1 - Arquitetura de segurança em camadas para sistemas de recomendação com IA.

Fonte: elaborado pelo autor, com base em Cavoukian (2010), McMahan *et al.* (2017) e Dwork *et al.* (2006).

5 ANÁLISE DA SHOPEE BRASIL

5.1 APRESENTAÇÃO DA PLATAFORMA E SEU CONTEXTO

A Shopee é uma plataforma de e-commerce de origem singaporiana, pertencente ao grupo Sea Limited, que passou a operar no Brasil em 2019. Em poucos anos, tornou-se uma das maiores plataformas de comércio eletrônico do país, com dezenas de milhões de usuários ativos e um catálogo de produtos que abrange praticamente todas as categorias de consumo. Sua expansão no mercado brasileiro foi acelerada por uma estratégia agressiva de preços, frete subsidiado e forte presença em dispositivos móveis.

Do ponto de vista regulatório, a Shopee opera no Brasil sob a égide da LGPD e, por meio de sua subsidiária de pagamentos (ShopeePay / SHPP Brasil), também está sujeita à regulamentação do Banco Central. Essa dupla sujeição regulatória torna a plataforma um caso de estudo particularmente rico para a análise de privacidade e segurança de dados.

5.2 ARQUITETURA DE RECOMENDAÇÃO DA SHOPEE

Embora a Shopee não divulgue publicamente os detalhes técnicos de sua arquitetura de recomendação, é possível caracterizá-la a partir de fontes secundárias e da análise de seu comportamento observável. O sistema opera como um algoritmo híbrido que combina múltiplas abordagens de personalização.

Os dados coletados para alimentar o sistema incluem histórico de navegação (produtos visualizados, pesquisados e adicionados ao carrinho), histórico de compras, categorias de interesse seguidas pelo usuário, informações demográficas fornecidas no cadastro (como localização e data de nascimento), avaliações publicadas, taxa de conversão por categoria e dados contextuais como horário de acesso e tipo de dispositivo utilizado.

Com base nesses insumos, o algoritmo personaliza continuamente múltiplas superfícies da plataforma: a página inicial de cada usuário, as seções 'Indicado para você' e 'Você pode gostar', os resultados de busca, as sugestões de itens complementares nas páginas de produto, e as comunicações por push notification e

e-mail. A plataforma também disponibiliza um sistema de anúncios (Shopee Ads) que utiliza os mesmos dados comportamentais para segmentar a exibição de produtos patrocinados.

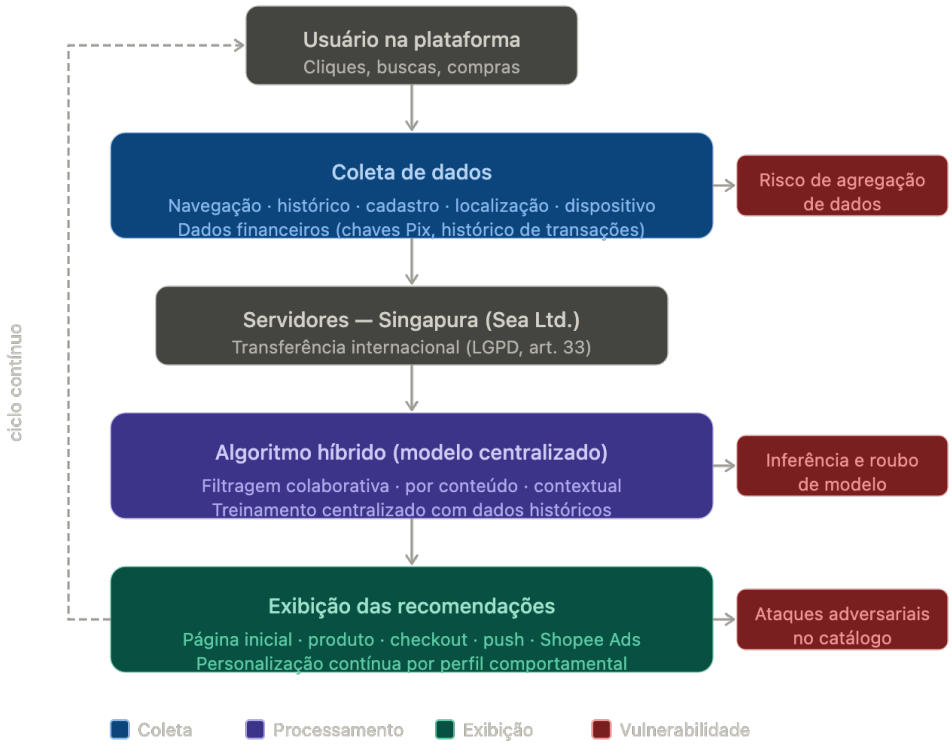
5.3 DADOS COLETADOS E SUPERFÍCIE DE EXPOSIÇÃO

A política de privacidade da Shopee Brasil declara a coleta de um amplo conjunto de dados pessoais dos usuários. Entre as categorias explicitamente mencionadas estão: dados de identificação e contato (nome, CPF, e-mail, telefone); dados financeiros (informações bancárias, chaves Pix, histórico de transações); dados de comportamento na plataforma (cliques, buscas, tempo de permanência por página, sequência de navegação); dados de localização; dados do dispositivo (modelo, sistema operacional, identificadores únicos); e dados derivados das avaliações e interações sociais na plataforma.

Do ponto de vista da segurança da informação, a combinação desses dados configura uma superfície de exposição significativa. Isoladamente, informações como 'usuário visualizou tênis de corrida às 22h de uma segunda-feira' parecem inofensivas. Contudo, quando agregadas ao longo do tempo e cruzadas com dados demográficos e financeiros, permitem inferir padrões comportamentais sensíveis, como condição financeira, rotinas domésticas, hábitos de saúde e até orientações de consumo que podem revelar aspectos da vida privada do usuário. Esse fenômeno ocorre devido a possibilidade de inferência a partir da agregação de dados comportamentais.

A política também esclarece que os dados coletados no Brasil podem ser transferidos para servidores localizados em Singapura, onde o grupo Sea Limited mantém sua infraestrutura central. Essa transferência internacional de dados está sujeita às obrigações do artigo 33 da LGPD, que exige que o país de destino ofereça grau de proteção adequado ou que sejam adotadas salvaguardas específicas.

Figura 2 — Fluxo de dados no sistema de recomendação da Shopee Brasil e vulnerabilidades identificadas.



Fonte: elaborado pelo autor, com base na política de privacidade da Shopee Brasil e em fontes secundárias.

5.4 VULNERABILIDADES IDENTIFICADAS NA SHOPEE BRASIL

A análise da arquitetura da Shopee à luz do referencial teórico apresentado nos capítulos anteriores permite identificar potenciais vulnerabilidades em cada categoria de ataque mapeada.

Em relação aos ataques de envenenamento, a Shopee opera um marketplace aberto em que qualquer vendedor pode criar uma conta e publicar produtos. Esse modelo cria condições propícias para a criação de perfis falsos de compradores visando manipular as métricas de relevância e posicionamento dos produtos no sistema de recomendação. A plataforma menciona em seus Termos de Serviço a proibição de 'realizar ações que possam burlar o sistema de feedback e avaliações',

o que confirma a ciência sobre esse vetor de ataque, mas a efetividade dos mecanismos de detecção não é auditável publicamente.

Quanto aos ataques de inferência, o sistema de recomendação da Shopee é treinado de forma centralizada, o que pode tornar suscetível a membership inference attacks. Um adversário com acesso à API pública da plataforma poderia, em tese, realizar consultas sistemáticas para inferir o comportamento de compra de usuários específicos. A ausência de evidências públicas de que a plataforma utilize privacidade diferencial no treinamento de seus modelos.

Do ponto de vista dos ataques de roubo de modelo, a Shopee expõe seu sistema de recomendação por meio de um aplicativo móvel com dezenas de milhões de instalações. O volume de consultas públicas disponíveis torna o sistema vulnerável a estratégias de extração por consultas sistemáticas, especialmente para as funcionalidades de busca e recomendação contextual.

Em relação aos ataques adversariais, o catálogo da Shopee abriga milhões de produtos cadastrados por vendedores externos. A ausência de uma camada de verificação adversarial das imagens e descrições de produtos significa que perturbações cuidadosamente calculadas poderiam ser inseridas no catálogo para manipular o comportamento do algoritmo de recomendação sem acionar os mecanismos de moderação existentes.

Uma análise crítica comparativa das vulnerabilidades identificadas permite hierarquizá-las por impacto e probabilidade. Os ataques de inferência representam o risco de maior impacto potencial, dado o volume de dados comportamentais acumulados pela plataforma: a exposição de padrões de compra sensíveis — medicamentos, itens de saúde, produtos de uso íntimo — configura risco direto de violação da privacidade de dados pessoais sensíveis nos termos do artigo 5º, II, da LGPD. Os ataques de envenenamento, por sua vez, representam a ameaça de maior facilidade de execução dado o modelo de marketplace aberto: a criação de perfis falsos de compradores não requer acesso privilegiado ao sistema. O roubo de modelo, embora tecnicamente mais sofisticado, tem motivação econômica clara em um mercado competitivo. A ausência de qualquer menção pública a mecanismos de privacidade diferencial ou aprendizado federado nos materiais técnicos e de conformidade da Shopee Brasil sugere que a plataforma opera com o modelo

centralizado tradicional, o que pode ampliar os riscos de inferência e minimizando a transparência sobre o tratamento dos dados de personalização.

5.5 INCIDENTE DE SEGURANÇA DE 2024

Em setembro de 2024, o Banco Central do Brasil comunicou publicamente a ocorrência de um incidente de segurança envolvendo dados pessoais vinculados a 150 chaves Pix sob responsabilidade da SHPP Brasil Instituição de Pagamento e Serviços de Pagamentos LTDA., subsidiária de pagamentos da Shopee. Segundo o Banco Central, o vazamento ocorreu entre os dias 2 e 4 de setembro de 2024 e foi causado por falhas pontuais nos sistemas da instituição.

As informações expostas eram de natureza cadastral: nome do usuário, CPF, instituição de relacionamento, agência, número e tipo da conta vinculados às chaves Pix afetadas. O Banco Central esclareceu que dados sensíveis sob sigilo bancário — como senhas, saldos e histórico de transações — não foram comprometidos.

A Shopee, em nota oficial, afirmou ter detectado “tentativa indevida para obter informações de chaves Pix” por meio de um recurso da plataforma que permitia aos usuários da ShopeePay visualizar informações dos destinatários de transferências. A empresa declarou ter implementado medidas de segurança adicionais imediatamente após tomar conhecimento do incidente.

Do ponto de vista analítico, o incidente ilustra com precisão o mecanismo de privacidade por agregação discutido neste trabalho: dados cadastrais como nome e CPF parecem inofensivos isoladamente, mas sua combinação com informações de conta bancária, em um contexto de e-commerce que já detém o histórico comportamental completo do usuário, cria um perfil de exposição consideravelmente mais amplo do que o incidente foi classificado oficialmente. O fato de o vazamento ter sido identificado não por mecanismos internos da plataforma, mas por um recurso de acesso público à API de consulta de destinatários, pode sugerir limitações nos controles de acesso e na monitoração de uso anômalo da plataforma.

5.6 CONFORMIDADE COM A LGPD

A Shopee Brasil demonstra conformidade formal com vários requisitos da LGPD. A plataforma mantém uma política de privacidade publicada e acessível, nomeia um Encarregado de Proteção de Dados (DPO), o contato

dpo.br@shopee.com está indicado publicamente, disponibiliza mecanismos para solicitação de acesso, correção e exclusão de dados, e informa sobre as hipóteses de compartilhamento com terceiros e transferências internacionais.

Contudo, há aspectos da conformidade que merecem questionamento crítico. A base legal utilizada para a maior parte do tratamento de dados é o consentimento, coletado por adesão aos Termos de Serviço. Essa prática tem sido questionada por especialistas em proteção de dados, pois o consentimento obtido como condição de acesso ao serviço pode não atender ao requisito de livre manifestação de vontade exigido pela LGPD. Adicionalmente, a política de privacidade não detalha os algoritmos de tomada de decisão automatizada utilizados no sistema de recomendação, o que dificulta o exercício do direito à explicação previsto no artigo 20 da LGPD.

O incidente de setembro de 2024 também levanta questões sobre a robustez dos processos de Avaliação de Impacto à Proteção de Dados (AIPD) realizados pela plataforma. Recursos que permitem acesso amplo a dados cadastrais de terceiros via API, como o recurso de consulta de destinatários de transferências que originou o vazamento, idealmente deveriam ser submetidos a avaliações rigorosas de risco antes de sua disponibilização.

5.7 LACUNAS E RECOMENDAÇÕES

A análise do caso Shopee Brasil permite identificar lacunas concretas entre as práticas observadas e as boas práticas recomendadas pela literatura e exigidas pela LGPD. Do ponto de vista técnico, as principais lacunas identificadas são: ausência de evidências de uso de privacidade diferencial no treinamento dos modelos de recomendação; centralização do treinamento dos algoritmos, que aumenta a superfície de ataque a *inference attacks*; ausência de mecanismos públicos de verificação adversarial das contribuições de conteúdo do catálogo; e controles insuficientes sobre o uso da API de consulta, como evidenciado pelo incidente de 2024.

As recomendações técnicas mais relevantes para mitigar essas lacunas são: adoção de privacidade diferencial no treinamento dos modelos de recomendação, começando com valores conservadores de ε e ajustando conforme o impacto na qualidade seja avaliado; implementação de aprendizado federado para as

funcionalidades de recomendação de maior sensibilidade; estabelecimento de limites de taxa e monitoramento de padrões anômalos de consulta à API para detectar tentativas de roubo de modelo; e validação adversarial automatizada das imagens e descrições inseridas por vendedores externos no catálogo.

Do ponto de vista regulatório, recomenda-se: revisão da base legal utilizada para o tratamento de dados de personalização, com maior clareza sobre o uso do legítimo interesse como hipótese alternativa ao consentimento; maior transparência sobre os critérios de decisão automatizada utilizados nos algoritmos de recomendação, em atendimento ao artigo 20 da LGPD; e condução de AIPDs periódicas para todos os recursos que envolvam acesso a dados de terceiros.

6 CONSIDERAÇÕES FINAIS

O percurso traçado ao longo desta pesquisa permitiu construir um panorama abrangente de um problema que tende a se tornar cada vez mais relevante: como garantir que os sistemas de recomendação com IA, centrais para a experiência do varejo digital, não se tornem instrumentos de vigilância ou vetores de vazamento de informações pessoais.

A revisão bibliográfica confirmou que os riscos não são hipotéticos. Os ataques de envenenamento, os métodos de inferência, os ataques adversariais e de roubo de modelo, e as vulnerabilidades de reidentificação têm base empírica sólida, documentada em experimentos controlados com ambientes próximos aos que as plataformas reais utilizam. A análise da Shopee Brasil evidenciou que esses riscos se materializam também em escala nacional: o incidente de setembro de 2024 demonstrou concretamente como falhas de controle de acesso em um sistema de e-commerce podem resultar em exposição de dados cadastrais de usuários.

O mapeamento dos mecanismos de proteção disponíveis mostrou que existem respostas técnicas para esses problemas. O aprendizado federado desacopla o treinamento de modelos da centralização de dados. A privacidade diferencial oferece garantias matemáticas de não rastreabilidade. O Privacy by Design garante que essas preocupações entrem na equação desde o início do desenvolvimento. A análise comparativa realizada no capítulo 4 evidenciou, contudo, que nenhum mecanismo isolado é suficiente: é a combinação deles que permite um equilíbrio favorável entre personalização e privacidade.

O marco regulatório, LGPD no Brasil e GDPR na Europa, representa um indutor importante de mudanças. A análise da Shopee Brasil revelou que a conformidade formal com a LGPD, embora presente em vários aspectos, não equivale à robustez técnica necessária para enfrentar as ameaças documentadas. A distância entre o que a lei exige e o que as plataformas efetivamente praticam permanece considerável, especialmente em relação à transparência dos algoritmos de decisão automatizada e à adoção de técnicas avançadas de proteção nos modelos de machine learning.

Este trabalho apresenta limitações que merecem reconhecimento. Por se tratar de pesquisa bibliográfica e documental complementada por estudo de caso com base em fontes públicas, não alcança a profundidade analítica que experimentos práticos

ou acesso a dados internos das plataformas permitiriam. Os sistemas de recomendação reais operam com arquiteturas proprietárias raramente detalhadas em publicações científicas, o que impede conclusões definitivas sobre implementações específicas.

Para pesquisas futuras, quatro direções mostram-se promissoras: (a) avaliações empíricas do impacto de diferentes valores do parâmetro ε da privacidade diferencial sobre a qualidade das recomendações em cenários reais de e-commerce brasileiro; (b) investigações sobre como plataformas de varejo digital brasileiras têm respondido às obrigações da LGPD na prática, por meio de estudos de caso com acesso a dados internos; (c) pesquisas sobre a percepção dos consumidores brasileiros acerca da troca entre personalização e privacidade; e (d) estudos sobre a efetividade dos mecanismos de detecção de ataques adversariais em catálogos abertos de marketplace.

A conclusão mais ampla que este trabalho permite extrair é que o desenvolvimento responsável de sistemas de recomendação com IA exige que as decisões técnicas e arquiteturais sejam tomadas com plena consciência de seus impactos sobre os direitos dos usuários. Dessa forma, a incorporação de mecanismos de privacidade e segurança deve ser tratada como parte integrante do desenvolvimento de sistemas de recomendação, contribuindo para maior confiabilidade, transparência e conformidade regulatória nas plataformas digitais.

REFERÊNCIAS

BERLIOZ, A. et al. Applying differential privacy to matrix factorization. In: **ACM CONFERENCE ON RECOMMENDER SYSTEMS**, 9., 2015, Vienna. Proceedings [...]. Vienna: ACM, 2015. p. 107-114.

BOBADILLA, J. et al. Recommender systems survey. **Knowledge-Based Systems**, Amsterdam, v. 46, p. 109-132, 2013.

Brasil. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Diário Oficial da União, Brasília, DF, 15 ago. 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 10 fev. 2025.

BURKE, R. Hybrid recommender systems: survey and experiments. **User Modeling and User-Adapted Interaction**, Dordrecht, v. 12, n. 4, p. 331-370, nov. 2002.

CADWALLADR, C.; GRAHAM-HARRISON, E. Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. **The Guardian**, London, 17 mar. 2018.

CHAI, D. et al. Secure federated matrix factorization. **IEEE Intelligent Systems**, Washington, v. 36, n. 5, p. 11-20, 2020.

CHEN, J. et al. Data poisoning attacks on factorization-based collaborative filtering. In: **ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS**, 2019, Vancouver. Proceedings [...]. Red Hook: Curran Associates, 2019. p. 1-12.

EUROPEAN PARLIAMENT; COUNCIL OF THE EUROPEAN UNION. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016. **Official Journal of the European Union**, Luxembourg, 4 maio 2016.

LOPS, P.; DE GEMMIS, M.; SEMERARO, G. Content-based recommender systems: state of the art and trends. In: RICCI, F. et al. (org.). Recommender systems handbook. New York: Springer, 2011. p. 73-105.

MCMAHAN, B. et al. Communication-efficient learning of deep networks from decentralized data. In: **INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND STATISTICS**, 20., 2017, Fort Lauderdale. Proceedings [...]. Fort Lauderdale: PMLR, 2017. p. 1273-1282.

NARAYANAN, A.; SHMATIKOV, V. Robust de-anonymization of large sparse datasets. In: **IEEE SYMPOSIUM ON SECURITY AND PRIVACY**, 29., 2008, Oakland. Proceedings [...]. Washington: IEEE, 2008. p. 111-125.

SHOPEE BRASIL. Política de Privacidade. Disponível em: https://help.shopee.com.br/portal/4/article/77068. Acesso em: mar. 2025.

SHOKRI, R. et al. Membership inference attacks against machine learning models. In: **IEEE SYMPOSIUM ON SECURITY AND PRIVACY**, 38., 2017, San Jose. Proceedings [...]. Washington: IEEE, 2017. p. 3-18.

GOODFELLOW, I. J. et al. Explaining and harnessing adversarial examples. In: **INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS,** 3., 2015, San Diego. Proceedings [...]. San Diego: ICLR, 2015.

PARK, S. et al. Privacy-preserving federated learning for recommender systems. In: **ACM SIGKDD CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING**, 27., 2021, Virtual. Proceedings [...]. New York: ACM, 2021. p. 1428-1437.

BIGGIO, B.; ROLI, F. Wild patterns: ten years after the rise of adversarial machine learning. Pattern Recognition, Amsterdam, v. 84, p. 317-331, 2018.

TRAMÉR, F. et al. Stealing machine learning models via prediction APIs. In: USENIX SECURITY SYMPOSIUM, 25., 2016, Austin. Proceedings [...]. Berkeley: USENIX Association, 2016. p. 601-618.

ZHANG, S. et al. Deep learning based recommender system: a survey and new perspectives. ACM Computing Surveys, New York, v. 52, n. 1, p. 1-38, 2019.

ZHANG, X.; CORMODE, G.; MIKLAU, G. Differentially private data publishing and analysis: a survey. IEEE Transactions on Knowledge and Data Engineering, Washington, v. 29, n. 8, p. 1619-1638, 2021.