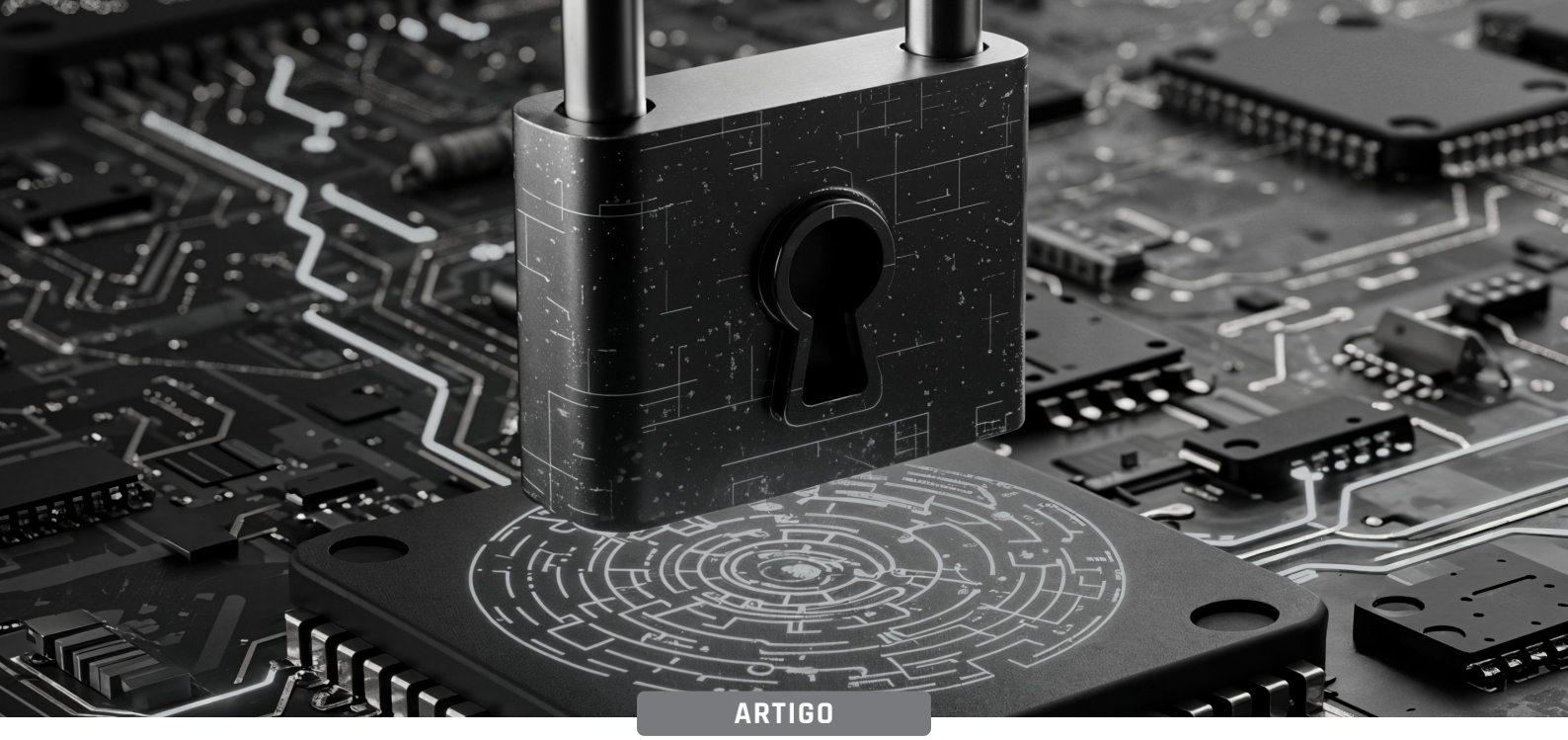

ARTIGO

CIBERSEGURANÇA FRENTE AOS RISCOS DA TRANSIÇÃO QUÂNTICA, DA INTELIGÊNCIA ARTIFICIAL, E DE ASPECTOS HUMANOS

POR
Marcos Antonio Simplicio Junior
msimplicio@usp.br

Os Referenciais de Formação do Curso de Bacharelado em Cibersegurança (RF-CS) construído pela Sociedade Brasileira de Computação [5] se apoia em oito eixos temáticos: Segurança de Dados, Segurança de Sistemas, Segurança de Conexão, Segurança de Software, Segurança de Componentes, Segurança Organizacional, Fatores Humanos em Segurança, e Segurança e Sociedade. Embora o objetivo primário do documento seja guiar a formação de profissionais de cibersegurança, ele também permite identificar tópicos amplos e seus respectivos desafios mais relevantes de pesquisa e inovação na área, os quais são descritos a seguir.

Preparar os sistemas computacionais para a transição quântica, com algoritmos e protocolos resistentes a ataques feitos via computadores quânticos.

Mitigar a ameaça da computação quântica está entre os principais desafios nos eixos de Segurança de Dados e Segurança de Conexões, que tratam da proteção de dados armazenados, em processamento, e enquanto atravessam interligações físicas e lógicas entre componentes computacionais. A razão é que, com computadores quânticos de grande porte, pode-se obter em tempo polinomial as chaves privadas dos principais esquemas criptográficos para assinatura digital e encapsulamento de chaves em

uso atualmente. Assim, embora aumentar a capacidade e confiabilidade de computadores quânticos ainda seja repleto de desafios, a sua superação pode trazer impactos graves à proteção de dados e comunicações em nível global.

As duas estratégias principais para a proteção de dados em um cenário pós-quântico são: (1) desenvolver esquemas criptográficos cuja segurança se apoie em problemas computacionais difíceis tanto em computadores clássicos como quânticos; e (2) para a transmissão de dados sigilosos, adotar esquemas quânticos de distribuição de chaves, explorando propriedades físicas do meio de comunicação. Essas linhas têm sido alvo de iniciativas de padronização, respectivamente, por órgãos internacionais como NIST [2] e ITU [1]. Apesar desse avanço, desafios importantes permanecem tanto para os esquemas já padronizados como para possíveis alternativas, considerando métricas de segurança, desempenho e flexibilidade. Em especial, destacam-se como áreas de pesquisa: (1) análise de segurança dos algoritmos e protocolos, cobrindo aspectos teóricos e práticos; (2) a otimização desses esquemas, considerando técnicas de projeto e de implementação, para que a eficiência resultante de sua adoção seja similar ou superior àquela fornecida por alternativas clássicas; e (3) a adaptação de soluções existentes para incorporar esquemas resistentes a computadores quânticos, preservando funcionalidades e a capacidade de adoção de técnicas modernas de projeto (por exemplo, arquiteturas de rede com uma camada de controle baseada em software).

No Brasil, iniciativas para sanar esse desafio, embora razoavelmente recentes, têm ganhado força nos últimos anos nos setores público e privado. Exemplos incluem o anúncio pelo Ministério da Ciência, Tecnologia e Inovação (MCTI), em abril/2025, da intenção de investir bilhões de reais em tecnologias quânticas, e também ações estratégicas em criptografia pós-quântica anunciadas com destaque por entidades como a Agência Brasileira de Inteligência (ABIN), o banco Bradesco, e a empresa de cibersegurança Kryptus.

Promover a evolução da Inteligência Artificial e da Ciência de Dados em um contexto de ameaças, mitigando riscos conhecidos e desconhecidos associados a essas tecnologias emergentes.

Tecnologias emergentes, fruto da Inovação e Empreendedorismo Tecnológico, costumam ser alvos bastante visados por atacantes, que buscam explorar a menor maturidade de segurança na integração entre seus componentes de hardware, software, comunicação e sub-sistemas. Essa preocupação, cerne do eixo de Segurança de Sistemas, e também associada aos eixos de Segurança de Software e de Componentes, atualmente tem se voltado a tecnologias emergentes como Inteligência Artificial (IA) e Ciência de Dados (CD). Em especial, um desafio notório com o crescente interesse em (e adoção de) mecanismos de IA refere-se à aplicação de Cibersegurança na construção e operação desses sistemas inteligentes, evitando que seus benefícios sejam ofuscados pelos riscos inerentes.

A pesquisa atual em torno desse desafio costuma ser associada a termos como Adversarial AI, ou "Aprendizado de Máquina Seguro" [6]. Esses termos complementam, assim, uma área de investigação mais tradicional, no sentido ``inverso", de aplicação de técnicas de IA e CD para aumentar a eficácia de ferramentas de segurança para monitoramento e detecção de ataques, identificação de vulnerabilidades em software, entre outras. Muito da literatura emergente na área consiste na construção de ferramentas de ataque, voltadas a identificar e explorar vulnerabilidades (como burlar filtros em sistemas de IA generativa, o chamado jailbreaking). Contudo, começam também a surgir propostas voltadas à mitigação de ataques de diferentes tipos: (1) perpetrados com o uso de IA, como a criação de deep fakes ou de malware para evadir proteções atuais; (2) direcionados a módulos de IA, visando ao roubo de dados ou modelos completos; ou (3) mirando sistemas que usam IA (por exemplo, explorando técnicas de jailbreaking, ou via envenenamento de bases de treinamento). Mitigar essas ameaças é desafiador devido à própria natureza das técnicas de aprendizado de máquina, em que ocorre a transformação de dados em programas. Nesse cenário, bastante distinto do desenvolvimento de software tradicional, é complexo definir políticas de segurança adequadas ou até mesmo seguir boas práticas comuns para proteção de sistemas.

No Brasil, aspectos de cibersegurança em IA são apresentados, ao menos em alto nível, como requisitos relevantes tanto no Plano Brasileiro de Inteligência Artificial (PBIA), lançado pelo MCTI em agosto/2024, quanto pelo Marco Regulatório da Inteligência Artificial (Projeto de Lei 2338/2023), que se encontrava em tramitação no momento da escrita deste documento. Ainda, essa intersecção entre cibersegurança e IA tem motivado o surgimento de centros de excelência em pesquisa em diferentes pontos do país, unindo esforços da academia, governo e setor privado. Dois exemplos ilustrativos, ambos lançados em 2025, são a Cátedra de IA Responsável, uma parceria entre Universidade de São Paulo (USP) e Google, e o Centro de Excelência em Inteligência Artificial para Segurança Cibernética, resultado de colaboração entre a Universidade Federal de Pernambuco (UFPE), Tempest, Embraer, Atech, e FAPESP.

Integrar e consolidar fatores humanos na cibersegurança: da educação dos usuários e desenvolvedores à usabilidade no projeto dos sistemas de proteção.

Os eixos de Segurança Organizacional, Segurança e Sociedade, e Fatores Humanos em Segurança têm um ponto em comum: a interação entre indivíduos e sistemas computacionais. Nesse contexto, o desafio computacional crítico que se impõe refere-se à necessidade de reduzir potenciais vulnerabilidades criadas dependendo das decisões e comportamentos adotados nessas interações, de forma intencional ou não [3]. Embora longe de ser um desafio novo, ele ainda é bastante presente na sociedade atual, um efeito direto da ubiquidade da computação e da Internet no mundo. Isso pode ser observado no crescente número de ata-

ques na cadeia de suprimentos de hardware/software por agentes maliciosos, como: a inserção de malware no software distribuído pela empresa Solarwinds em 2019 e na biblioteca XZ Utils em 2024; o acoplamento de explosivos em dispositivos de comunicação adquiridos pelo Hezbollah em 2024; e a fraude via recrutamento de agente interno na C&M em 2025. Já casos originados em falha não intencional são observados na desastrosa atualização de software realizada pela CrowdStrike em 2024, que causou um apagão mundial de sistemas, e também em casos de engenharia social que, segundo a Febraban, foram responsáveis por 70% das fraudes bancárias em 2023.

Pesquisas voltadas a sanar esses desafios costumam envolver duas abordagens principais, complementares. A primeira delas busca transformar os humanos, tradicionais elos fracos na proteção de sistemas, em aliados do processo de cibersegurança. Um exemplo são propostas voltadas à educação de usuários, seja por meio de: treinamento formal, de aspecto geral (por exemplo, seguindo o RF-CS) ou com foco em áreas críticas; e campanhas de conscientização eficientes, abordagem essencial para evitar falhas, ataques de engenharia social, e também para ajudar no combate à desinformação. Outro exemplo são soluções de "segurança com usabilidade" (usable security), estratégia que busca tornar sistemas de cibersegurança mais amigáveis e personalizáveis para usuários de forma geral, aplicando conceitos de áreas como psicologia, ciências sociais e interação humano-computador (IHC). Já a segunda abordagem busca incrementar a automatização e granularidade dos mecanismos de segurança, de forma invisível aos usuários. Soluções modernas nessa linha costumam seguir o conceito conhecido como Confiança Zero [4]: apoiando-se em sistemas de gestão de identidade, autenticação/autorização, e detecção/prevenção de intrusões, as interações entre módulos de um sistema ou entre sistemas são constantemente validadas. Ao mesmo tempo, empregam-se políticas de segurança adaptáveis, baseadas em riscos, seguindo-se boas práticas como privilégio mínimo e separação de deveres para limitar eventuais danos em casos de violação de segurança.

No Brasil, este tema é abordado com algum destaque na Estratégia Nacional de Cibersegurança (E-Ciber), lançada em 2025 por meio do Decreto nº 12.573, e também alvo de preocupação por setores particularmente afetados, como o financeiro e de infraestrutura. Ele também está no cerne de iniciativas voltadas à formação de profissionais especializados em cibersegurança, como o próprio RF-CS e o Programa "Hackers do Bem" – promovido pelo MCTI, com pela Rede Nacional de Ensino e Pesquisa (RNP), Senai-SP, e Softex.

Referências

1. ITU-T (2019). Recommendation ITU-T Y.3800 - overview on networks supporting quantum key distribution. Relatório técnico, International Telecommunication Union.
2. NIST (2023). SP 1800-38B: Migration to Post-Quantum Cryptography. Quantum Readiness: Cryptographic Discovery. National Institute of Standards and Technology.
3. Pollini, A., Callari, T. C., Tedeschi, A., Ruscio, D., Save, L., Chiarugi, F. e Guerri, D. (2022). Leveraging human factors in cybersecurity: an integrated methodological approach. Cognition, Technology & Work, 24(2):371–390.
4. Rose, S., Borchert, O., Mitchell, S. e Connelly, S. (2020). NIST SP 800-207: Zero trust architecture. Relatório técnico, National Institute of Standards and Technology.
5. SBC (2023). Referenciais de formação para o curso de bacharelado em cibersegurança. Relatório técnico, Sociedade Brasileira de Computação (SBC).
6. Vassilev, A., Oprea, A., Fordyce, A. e Anderson, H. (2024). Adversarial machine learning: A taxonomy and terminology of attacks and mitigations. Relatório técnico, National Institute of Standards and Technology.

.

MARCOS ANTONIO SIMPLICIO JUNIOR é Professor Associado na Escola Politécnica da Universidade de São Paulo (EP-USP). Atua em projetos de pesquisa relacionados a criptografia e cibersegurança desde 2007, comumente em parceria com os setores público e privado. Sua pesquisa resultou em mais de uma centena de trabalhos publicados, incluindo artigos acadêmicos, patentes e contribuições para padrões internacionais. É atualmente o coordenador da Comissão Especial de Cibersegurança da Sociedade Brasileira de Computação (CESeg-SBC) e do Laboratório de Arquitetura e Redes de Computadores (LARC) da USP.