

# **SOB ATAQUE!** **A REALIDADE INVISÍVEL DAS AMEAÇAS** **CIBERNÉTICAS EM IES**



**FÁBIO XAVIER**



# +30 ANOS

## EXPERIÊNCIA PROFISSIONAL

# +22 ANOS

## na academia

# Colunista

MIT  
Technology  
Review

Publicado por TEC



itforum



## Fábio Correa Xavier

@fabio@tce.sp.gov.br

https://www.linkedin.com/in/fabiocorreaxavier

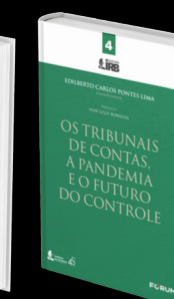
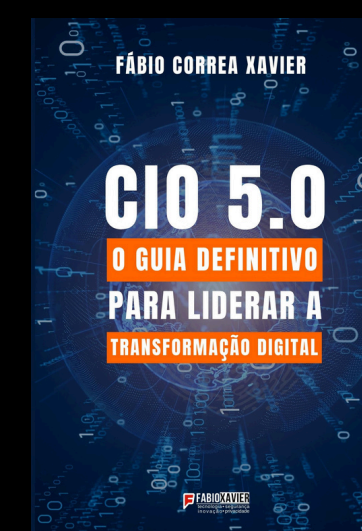
@fabiocx

fabioxavier.com.br



bits.fabioxavier.com.br

Assine o BITS



# ATO 1

# O IMPACTO INVISÍVEL

CYBER

ATTACK

**VOCÊ CONFIARIA SUA IES AO  
ACASO DIGITAL?**

# O Cibercrime Custará Ao Mundo US\$ 10,5 Trilhões Anualmente Até 2025



*Relatório especial: Guerra cibernética na alta gerência.*

- [Steve Morgan](#), Editor-chefe

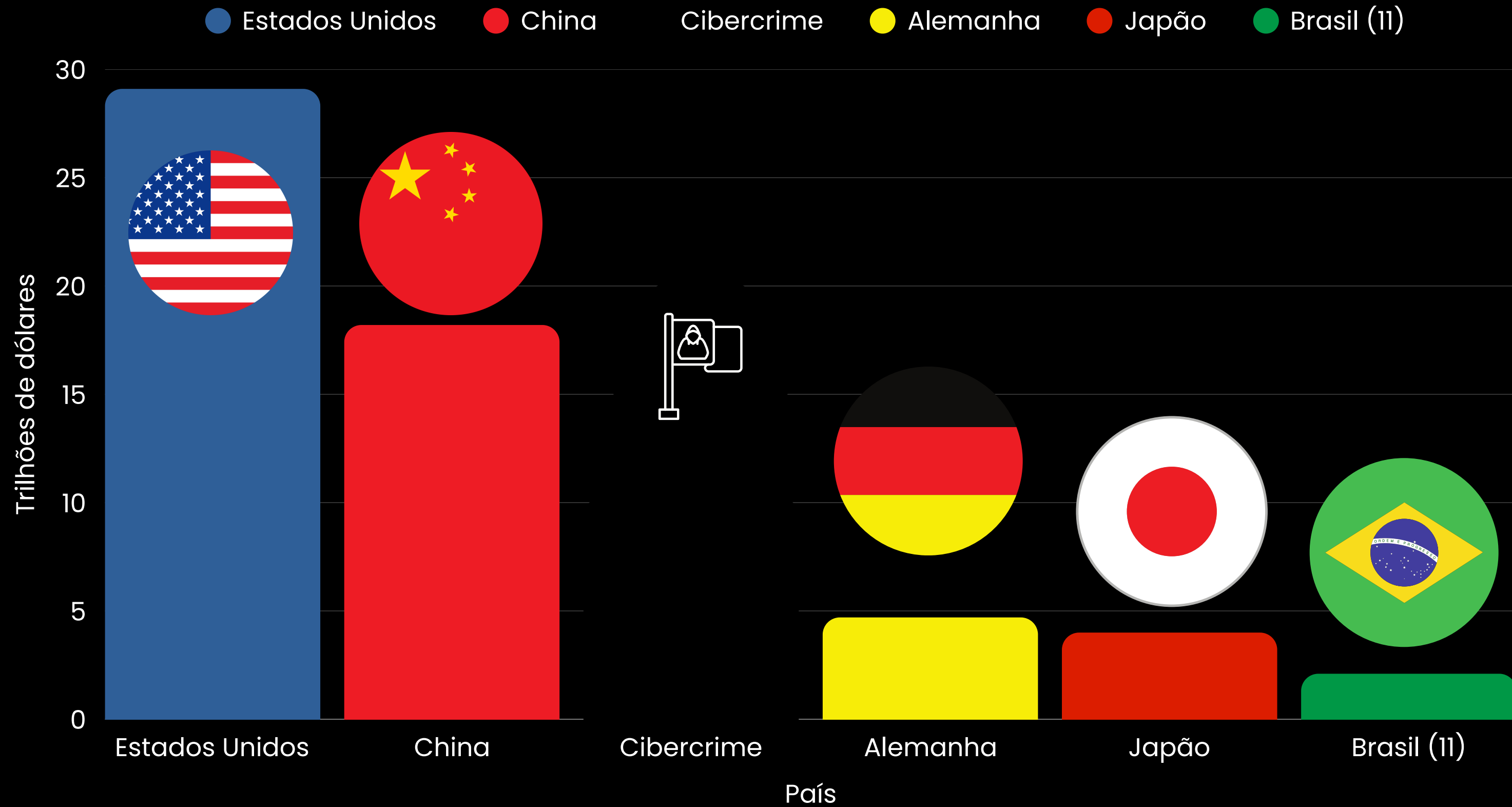
Sausalito, Califórnia - 13 de novembro de 2020

Se fosse medido como um país, o crime cibernético — que deve causar danos totalizando US\$ 6 trilhões em todo o mundo em 2021 — seria a terceira maior economia do mundo, depois dos EUA e da China.

A Cybersecurity Ventures prevê que os custos globais com crimes cibernéticos crescerão 15% ao ano nos próximos cinco anos, atingindo US\$ 10,5 trilhões anualmente até 2025, acima dos US\$ 3 trilhões em 2015. Isso representa a maior transferência de riqueza econômica da história, coloca em risco os incentivos para inovação e investimento, é exponencialmente maior do que os danos causados por desastres naturais em um ano e será mais lucrativo do que o comércio global de todas as principais drogas ilegais combinadas.



# PIB 2024



Nacional

## Ataque hacker: PF prende 2 suspeitos de participação em desvio de R\$ 541 mi

Grupo transformou dinheiro desviado em criptoativos; prisões foram em Goiás

Elijonas Maia, Thomaz Coelho, da CNN, em Brasília e São Paulo

16/07/25 às 17:52 | Atualizado 17/07/25 às 11:14

Ataque hacker: PF prende 2 suspeitos de participação em desvio de R\$ 541 milhões | CNN NOVO DIA



**PF PRENDE  
SUSPEITOS POR  
MAIOR ATAQUE  
HACKER DO BRASIL**



Nacional

## PF prende hackers que invadiram sites de tribunais de Justiça

Operação foi realizada em São Paulo, no Paraná e no Distrito Federal; sites do STF, STJ, CNJ e do Ae Guarulhos foram alvos de ataques

Elijonas Maia, da CNN, em Brasília

10/06/25 às 08:59 | Atualizado 10/06/25 às 09:19



## Hacker de 18 anos é preso por invasão ao TJ-PI

Jovem teria emitido mandados de prisão e realizado bloqueios bancários contra outro hacker.

08 de agosto de 2025 - 13:22

⚠ Tem algo errado neste post? Avise a redação!

Tamanho da fonte: -A +A



Tribunal de Justiça do Piauí - Foto: Site do TJ-PI

Um jovem de 18 anos, suspeito de um ataque hacker ao Tribunal de Justiça do Piauí (TJ-PI), foi preso na cidade de Itumbiara, no interior de Goiás.

# Educação é o 3º setor mais visado

**2.507** tentativas semanais de ataque em IES

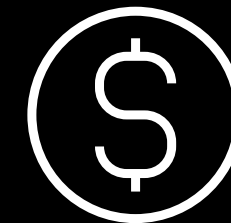
# POR QUE A EDUCAÇÃO É UM ALVO?



**Dados**  
sensíveis de  
milhões



**Ambiente**  
Cultura de  
abertura e  
colaboração



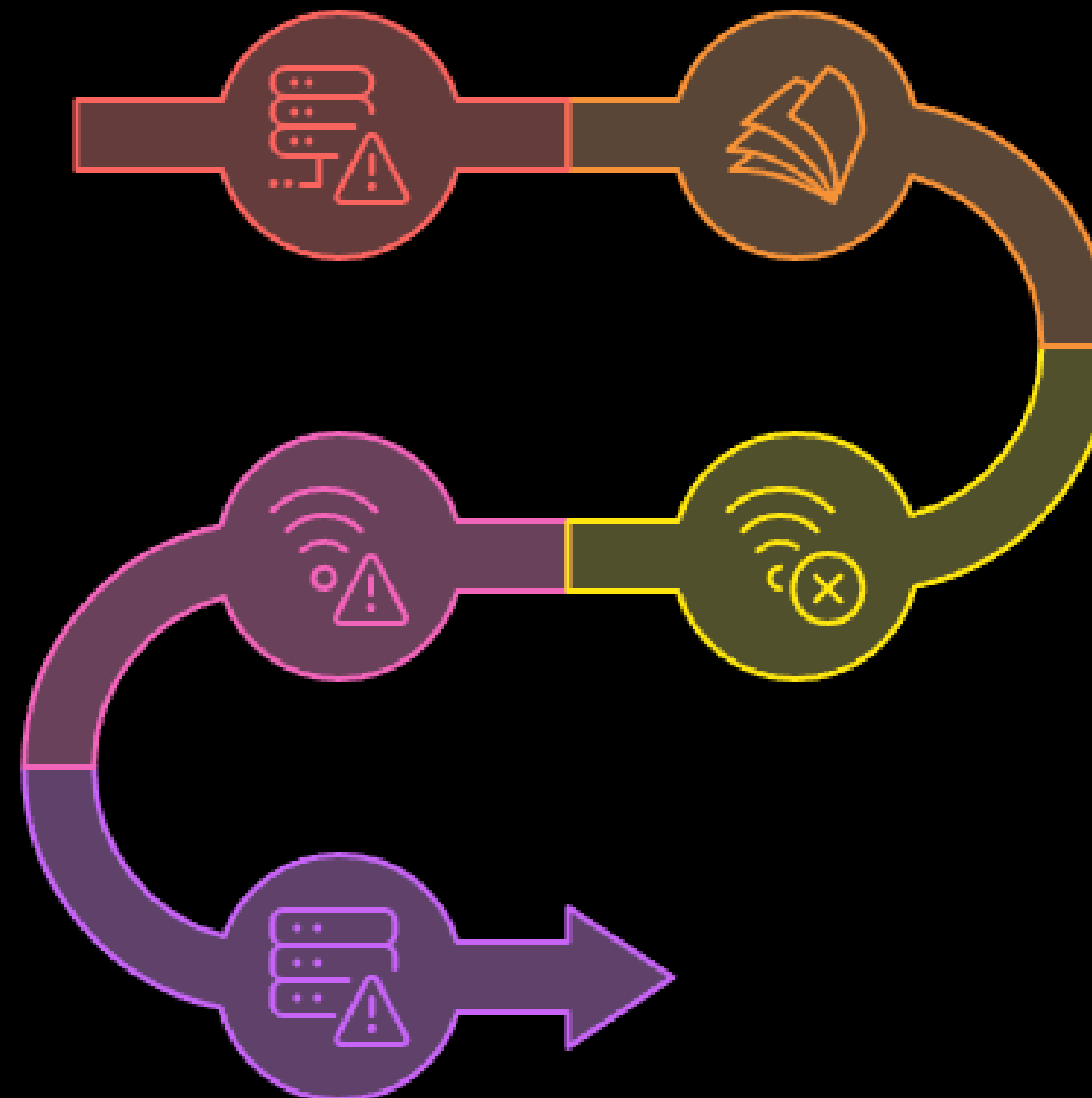
**20%**  
com  
orçamento  
adequado

# Linha do tempo de incidentes de cibersegurança em IES 2024-2025

**23/jul/2024**  
Ataque DDoS interrompe  
site e serviços

**10/abr/2025**  
Falha de firewall e  
tentativa de intrusão

**16/jul/2025**  
Ataque de ransomware  
bloqueia servidores



**11-27/mar/2025**  
Instabilidade nos  
sistemas acadêmicos

**09/abr/2025**  
Interrupção do portal e  
Wi-Fi

# RANSOMWARE 2025

**+130**

ataques no 1º  
semestre

**+23%**

Em relação a  
2024

**556k**

Resgate médio,  
em US\$

# RANSOMWARE 2025

**8%**

Recuperação  
total

**2.8M**

Custo médio, em  
US\$, por incidente

# IMPACTO INSTITUCIONAL

- #1** INTERRUPÇÃO DE AULAS E SERVIÇOS
- #2** PERDA DE PESQUISAS
- #3** REPUTAÇÃO ABALADA
- #4** PERDAS FINANCEIRAS
- #5** ANPD

**+75%**

**aumento dos ataques desde 2019**

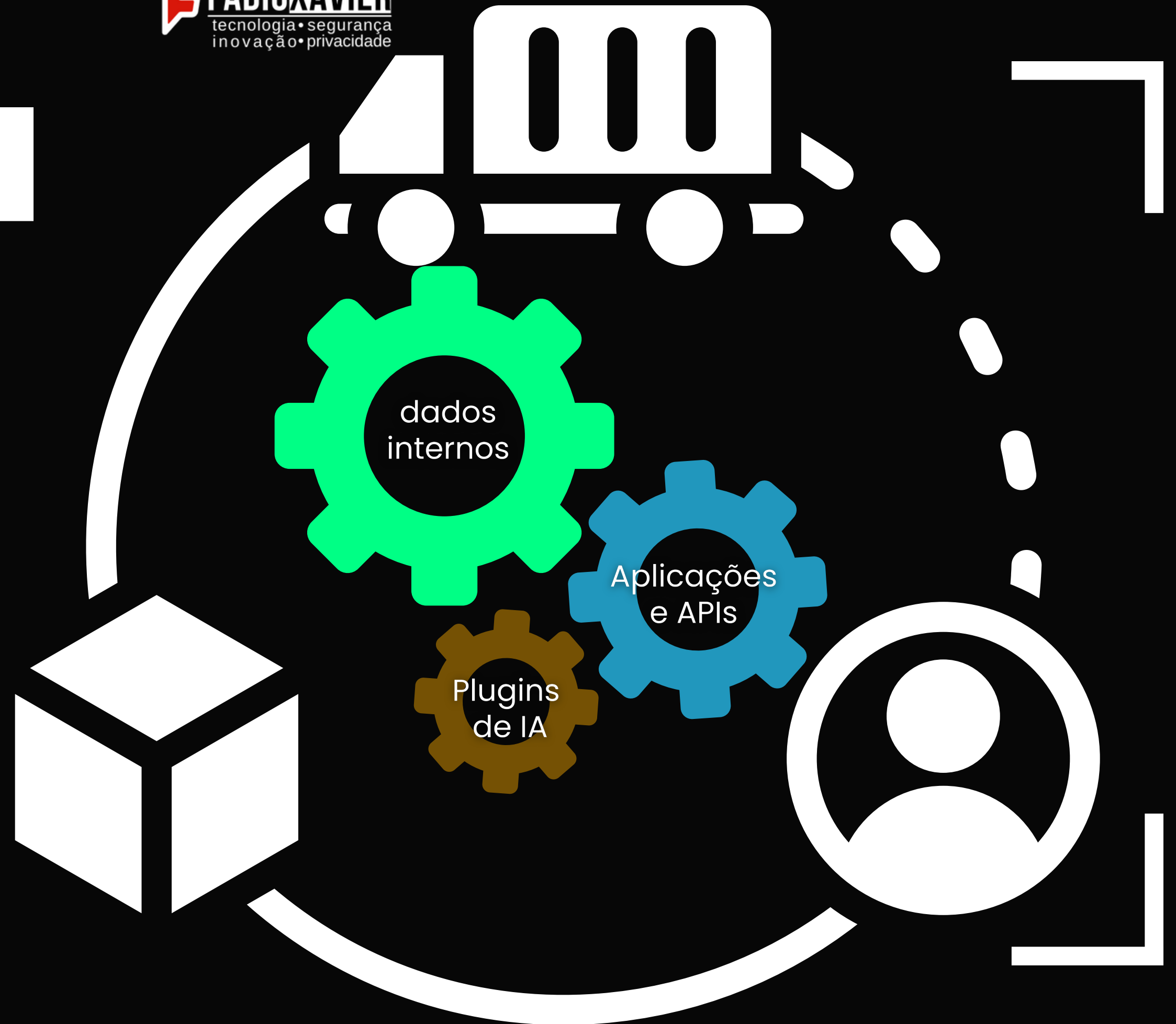
**ATO 2**

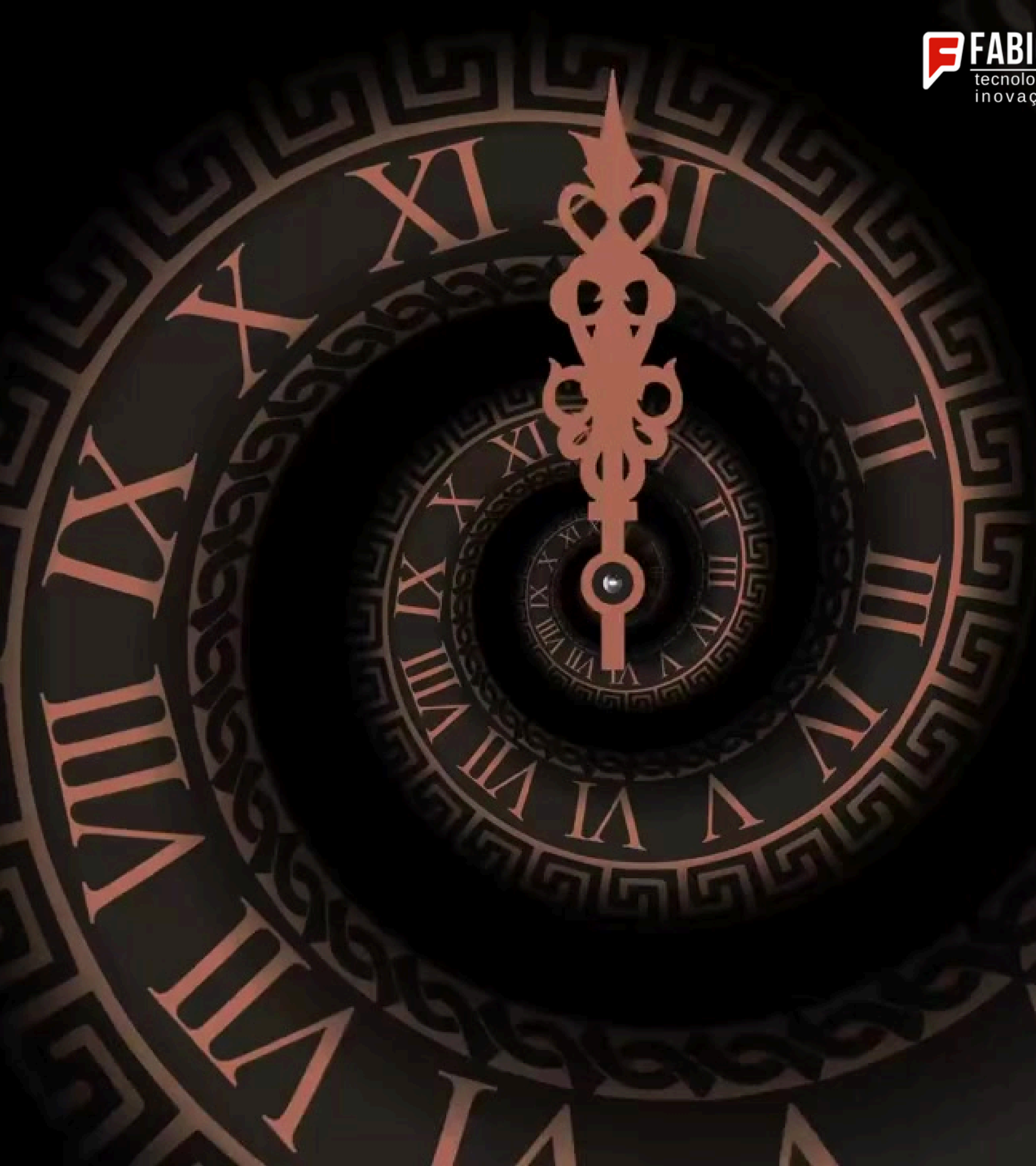
# A ESCALADA COM IA

**DO INVISÍVEL AO INEVITÁVEL**

# O QUE MUDOU COM A IA?

- + Superfície de ataque
- + velocidade de ataque
- + Shadow AI e Deepfake





```
char grade = 'A';
int age = 23;
float grade = 3.8;
double money = 1556.3204843;

main()
{
    // do some thing
}

do some thing
main()
{
    printf("Hello World");
}

int money = 100;
printf("money = %d", money);
money + 22;

char grade = 'A';
printf("Money = %d and Grade = %c", money, grade);

#include <stdio.h>
main()
{
    char alphabet_A = 'A';
    char alphabet_B = 'B';
    char alphabet_C = 'C';
    printf("\n A = %d", alphabet_A);
    printf("\n B = %d", alphabet_B);
}

printf("Hello World");
// do some thing
}

int value_A = 55;
int value_Z = value_A + 25;
printf("\n A+25 = %c", value_Z);

#include <stdio.h>
money = 100;
```

# ENGENHARIA SOCIAL COM IA

**+15.000**

QR Codes maliciosos/dia no setor  
educacional

**60% DAS VIOLAÇÕES ENVOLVEM  
PESSOAS.**

# SHADOW AI



## SUPERFÍCIE DE ATAQUE AMPLIADA

**244**

Domínios em  
média por IES

**10%**

Com portas RDP  
abertas

**48%**

Com softwares  
vulneráveis

**CUSTO MÉDIO DA VIOLAÇÃO**

**US\$ 4,4 MILHÕES (2025)**

**FATOR HUMANO: ~60% DAS VIOLAÇÕES**

**US\$ 7,19 MILHÕES (2025)**

**15 BILHÕES DE TENTATIVAS DE ATAQUES (1º SEMESTRE 2025)**

## BRASIL NO EPICENTRO

# 269 mil

incidentes jan-jul de 2025

# BRASIL NO EPICENTRO

**28,9%**

das origens de ataques globais

**ATO 3**

**A DEFESA  
NECESSÁRIA**

**CONSTRUINDO NOSSA  
DEFESA**

**28,9%**

**das origens de ataques globais**

## CONSTRUINDO NOSSA DEFESA - O MÍNIMO

 **99%** MFA reduz 99% dos riscos

 **75%** Treinamentos reduzem phishing em 75%

## BACKUPS 3-2-1

- 3 cópias dos dados
  - produção + 2 backups
- 2 tipos diferentes de mídia
  - nuvem e local, por exemplo
- 1 cópia off-site
  - fora do campus ou em nuvem segregada

# ARQUITETURA ZERO TRUST



## 1. Princípios

- Nunca confiar
- Verificar sempre
- Privilégio mínimo

## 2. Exemplos

- MFA
- Segmentação
- Microperímetro
- Perímetro expandido

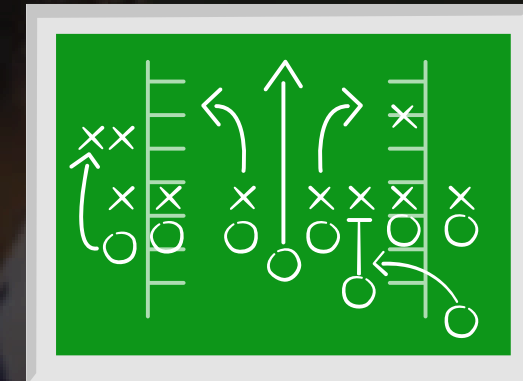
# PESSOAS E PROCESSOS PRIMEIRO



Treinamento  
contínuo  
anti-phishing



Simulações  
mensais



Playbooks de  
resposta

**FUNDAMENTOS**

+

**GOVERNANÇA DE IA**

+

**MÉTRICAS**

**Fundamentos**

**Governança de IA**

**Disciplina operacional  
orientada por métricas**

30 dias

Treinamentos

Avaliar modelos

Definir Processos

60 dias

Padronizar dados

Monitorar decisões

Estabelecer KPIs

90 dias

Análises contínuas

Auditoria externa

Melhoria contínua

# TOP 10 RISCOS DE LLMS E GEN AI | 2025

**LLM01: 2025**  
**Prompt Injection**

**LLM01:2025 Prompt Injection**

A Prompt Injection Vulnerability occurs when user prompts alter the...

[Read More](#)

**LLM02: 2025**  
**Sensitive Information Disclosure**

**LLM02:2025 Sensitive Information Disclosure**

Sensitive information can affect both the LLM and its application...

[Read More](#)

**LLM03: 2025**  
**Supply Chain**

**LLM03:2025 Supply Chain**

LLM supply chains are susceptible to various vulnerabilities, which can...

[Read More](#)

**LLM04: 2025**  
**Data and Model Poisoning**

**LLM04:2025 Data and Model Poisoning**

Data poisoning occurs when pre-training, fine-tuning, or embedding data is...

[Read More](#)

**LLM05: 2025**  
**Improper Output Handling**

**LLM05:2025 Improper Output Handling**

Improper Output Handling refers specifically to insufficient validation, sanitization, and...

[Read More](#)

**LLM06: 2025**  
**Excessive Agency**

**LLM06:2025 Excessive Agency**

An LLM-based system is often granted a degree of agency...

[Read More](#)

**LLM07: 2025**  
**System Prompt Leakage**

**LLM07:2025 System Prompt Leakage**

The system prompt leakage vulnerability in LLMs refers to the...

[Read More](#)

**LLM08: 2025**  
**Vector and Embedding Weaknesses**

**LLM08:2025 Vector and Embedding Weaknesses**

Vectors and embeddings vulnerabilities present significant security risks in systems...

[Read More](#)

**LLM09: 2025**  
**Misinformation**

**LLM09:2025 Misinformation**

Misinformation from LLMs poses a core vulnerability for applications relying...

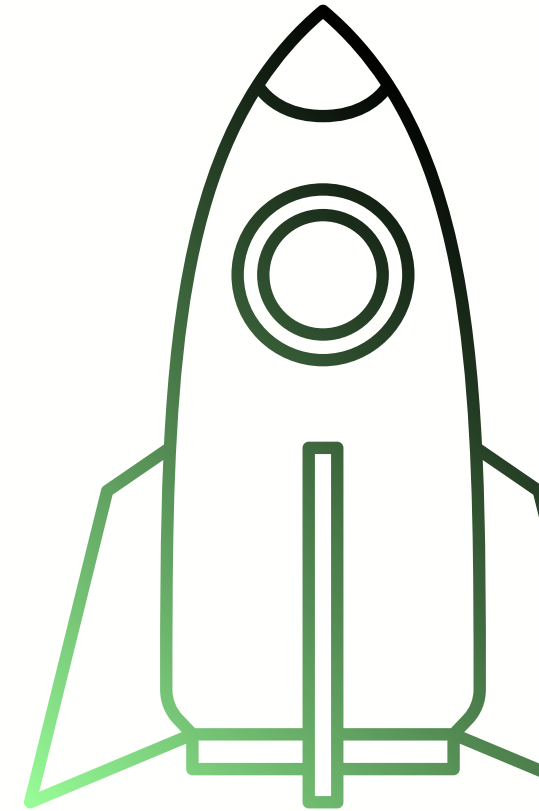
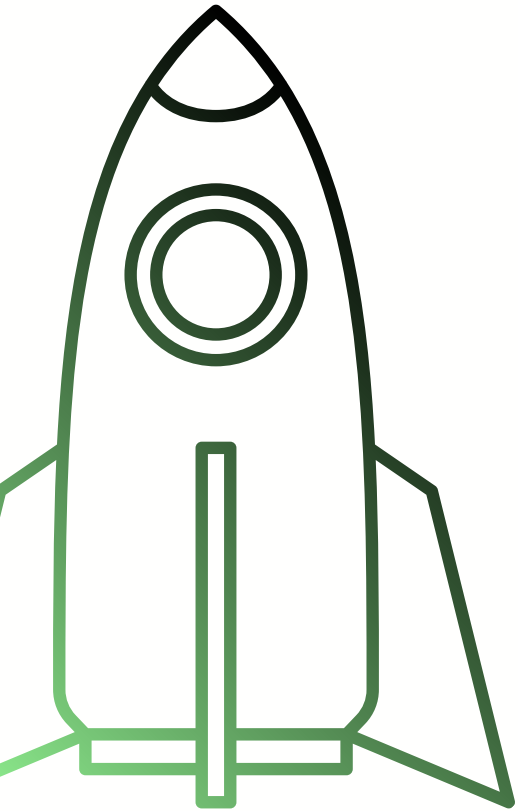
[Read More](#)

**LLM10: 2025**  
**Unbounded Consumption**

**LLM10:2025 Unbounded Consumption**

Unbounded Consumption refers to the process where a Large Language...

[Read More](#)



# VULNERABILIDADES TÉCNICAS

01

## Prompt Injection

Exploração de instruções inseridas para manipular a saída do modelo.

02

## Sensitive Information Disclosure

Exposição de dados sensíveis.

03

## Supply Chain

Riscos na cadeia de fornecimento de modelos e dados.

04

## Data and Model Poisoning

Manipulação maliciosa de dados de treino.

05

## Improper Output Handling

Falhas ao lidar com respostas incorretas ou não filtradas.

# RISCOS OPERACIONAIS E DE USO

01

## Excessive Agency

Delegação de ações além do controle esperado.

02

## System Prompt Leakage

Exposição de prompts internos.

03

## Vector and Embedding Weaknesses

Vulnerabilidades em vetores e embeddings.

04

## Misinformation

Geração e propagação de desinformação.

05

## Unbounded Consumption

Uso excessivo ou não monitorado de recursos.

# 5 PRINCÍPIOS DE DEFESA DA IA

## 1. Limitar Poderes

Aplicar o princípio do menor privilégio, controlar acessos e exigir aprovação humana em ações críticas.

## 2. Filtrar e validar

Sanitizar entradas e saídas, definir formatos claros de resposta e validar outputs continuamente.

## 3. Proteger informações

Isolar e criptografar prompts internos, mascarar dados sensíveis e monitorar acessos e logs.

## 4. Garantir integridade dos dados

Avaliar fornecedores, validar datasets, usar múltiplas fontes e realizar testes adversariais.

## 5. Gerir consumo e confiança

Definir quotas de uso, monitorar recursos, aplicar mecanismos de fallback e checar informações com fontes confiáveis.

# DEFESA EM CAMADAS

1

**TECNOLOGIA**

2

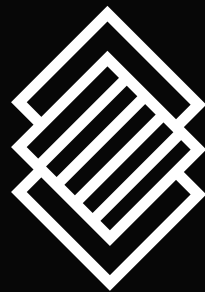
**PROCESSOS**

3

**CULTURA**

# **ATO FINAL (OU INICIAL) DA CONSCIÊNCIA À DECISÃO**

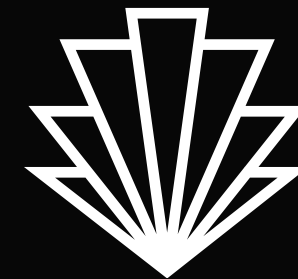
# OS PRÓXIMOS PASSOS



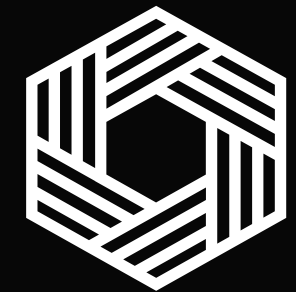
**IMPLEMENTAR MFA**



**PLANO DE BACKUP  
TESTADO**



**PROGRAMA DE  
CONSCIENTIZAÇÃO**



**REDUÇÃO DA  
SUPERFÍCIE**

**O FUTURO SEGURO COMEÇA  
HOJE**

**SEGURANÇA COMO MISSÃO EDUCACIONAL**

# CULTURA DE CORRESPONSABILIDADE

**DIRIGENTES, PROFESSORES, ALUNOS.**



**NÃO PODEMOS  
ENSINAR O FUTURO  
SE NÃO  
PROTEGERMOS O  
PRESENTE.**

**CIBERSEGURANÇA NÃO É SÓ TÉCNICA  
É GOVERNANÇA.**

**INCLUIR ESPECIALISTAS EM INOVAÇÃO  
E CIBERSEGURANÇA NO CONSELHO DE  
ADMINISTRAÇÃO DAS IES.**

**ONDE AS DECISÕES ESTRATÉGICAS SÃO TOMADAS,  
A SEGURANÇA PRECISA ESTAR PRESENTE.**

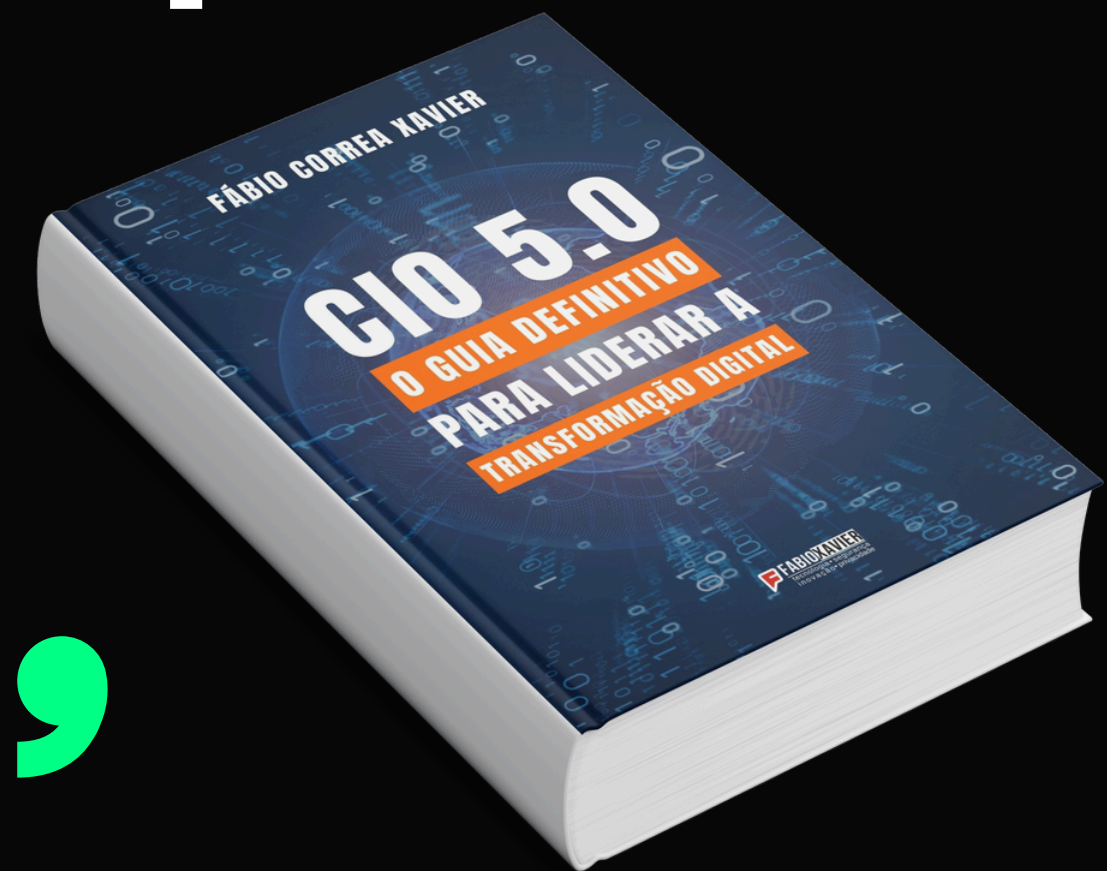


<https://fabioxavier.com.br/go/bits44>

“

**A SEGURANÇA CIBERNÉTICA NÃO É APENAS  
UMA QUESTÃO TÉCNICA, MAS ESTRATÉGICA,  
EXIGINDO UM ALINHAMENTO DIRETO COM A  
ALTA GESTÃO DA ORGANIZAÇÃO.**

”



---

# OBRIGADO!

---



**Fábio Correa Xavier**

[fabioxavier.com.br](http://fabioxavier.com.br)

Assine o BITS



[bits.fabioxavier.com.br](http://bits.fabioxavier.com.br)

