Compliance Killed Your Curiosity

Why Fear-Based Security Training Makes You Worse at Your Job

João Pedro Dias

https://jpdias.me

SEPRJ - ISEP

24/04/2026

whoami

Senior Software Architect @Kuehne+Nagel

PhD in Informatics Engineering @Univ. Porto

Jack of all trades, master of some.

jpdias@pm.me // jpdias.pt@gmail.com // https://www.linkedin.com/in/joaopdias/

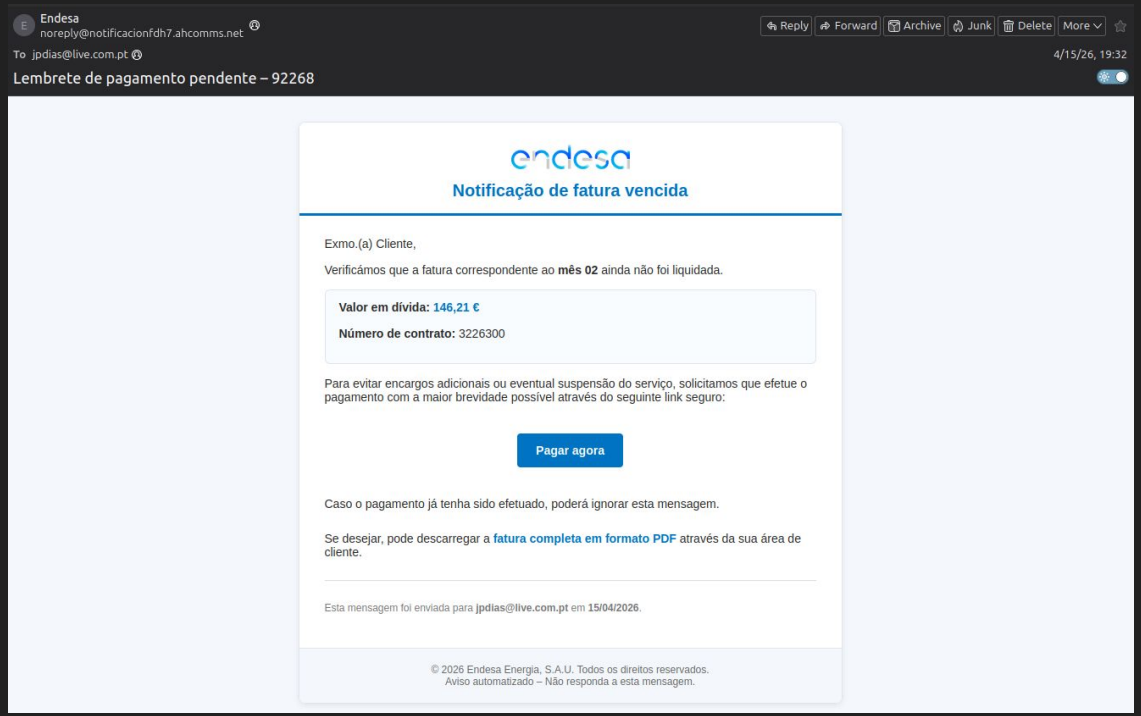
Disclaimer

Opinions expressed are solely my own and do not express the views or opinions of my employer.

Security training is making you (and your company*) worse at security.

* *Understanding the Efficacy of Phishing Training in Practice, https://www.computer.org/csdl/proceedings-article/sp/2025/223600a076/21B7RjYyG9q*

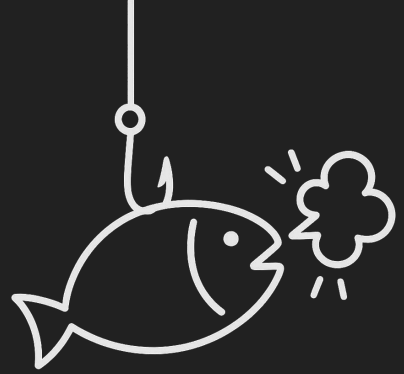
What do you do when you get a phishing email?



- ☐ Ignore
- ☐ Delete
- ☐ Report
- ☐ Click! (*)

Story time zzz…

Two months ago, I got *yet another* phishing email.



Standard training says: delete it.

Standard habit says: ignore…

http://something something

I clicked it instead.

And with it, according to every security training I've taken, I did everything wrong.

And that's when I learned more than any security training ever taught me.

3.4B phishing emails are sent globally every single day[1]

Attackers learn from every email (they test, adapt, improve).

You ignore/delete every email.

This is **asymmetric learning**.

Who gets better over time?

[1] *Phishing statistics 2025 – 2026: The numbers you need to know, https://zensec.co.uk/blog/2025-phishing-statistics-the-alarming-rise-in-attacks/*

This isn’t just asymmetry.
It’s accelerating asymmetry.

AI = Asymmetric Scaling

AI is scaling phishing. Unevenly.

Attackers now use AI to:

- Generate better lures
- Personalize at scale
- Iterate faster than ever

Result:

- More campaigns
- Faster evolution
- New techniques appear constantly

We need more people looking, analyzing, and learning — not fewer clicks.

What about Spam Filters?

Email authentication can be bypassed:

- SPF/DKIM/DMARC: Attackers use legitimate services (Google Drive, Azure, SendGrid, Twilio, etc.)
- Trusted senders: Compromised accounts or SMTP servers = trusted headers
- Evilginx/reverse proxies: Real-time MitM of 2FA sessions
- Header spoofing: Display name ≠ actual sender

Filters catch volume (and lazy) attacks. Sophisticated campaigns slip through.

Call for Action

Why did we stop being curious?

Cover Your Ass (CYA) culture

"We trained everyone not to click" is a legal defense, not *effective* security.

ISO 27001 (and the like) protects organizations from lawsuits…

…not you from staying mediocre and disinterested.

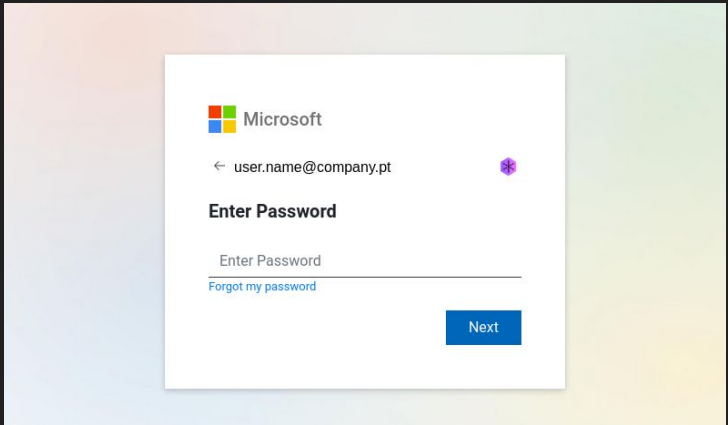
You *are* allowed to investigate

- ✓ Analyzing public URLs (no auth required)
- ✓ Running samples in sandboxes (your environment)
- ✓ WHOIS lookups (public data)
- ✓ Abuse reporting (legitimate channel)

No permission need, this is normal security research.
No different than: Reverse engineering, vulnerability research, CTFs, etc.

Motivational Examples

On the hook of a phisher (2022)



Phishing for stealing MS credentials and two-factor codes. They used Clearbit API to fake legitimacy.

Tools: URLScan.io and JSNice

* *On the hook of a phisher, https://jpdias.me/infosec/2022/07/30/phishing-story.html*

Plywood Trojan (2025)

Budget crypto wallet phishing using RMM tool (similar to AnyDesk), which is not flagged by most antivirus/EDR solutions.

Tools: URLScan.io, Claude for code analysis

If they can do this on a budget (with AI), you can investigate on a budget.

* *Plywood Trojan: When Attackers Go Budget, https://jpdias.me/infosec/2025/11/17/trojans-made-easy.html*

VoicePress5 (2026)

FATURA

Número da fatura:	900558819
Data da fatura:	28/01/2026
Estado:	Em aberto (não paga)

Ver fatura

Esta fatura encontra-se atualmente por liquidar.

Com o objetivo de garantir a máxima transparência na faturação, caso tenha alguma dúvida poderá contactar-nos por e-mail.

Solicitamos o envio do comprovativo de pagamento após a respetiva liquidação.

Phishing with multi-stages for spreading a Java RAT (Ratty). The RAT was re-written (i.e. AI) thus bypassing most antivirus signature checks.

Tools: URLScan.io, Claude for code analysis, ANY.RUN for behaviour analysis

* *VoicePress5: Tracing a Phishing-to-Java RAT Infection Chain, https://jpdias.me/infosec/2026/03/04/email-to-rat.html*

Passive vs. Active Threat Model

Passive defense:

- Assume: Threat intelligence comes from vendors
- Result: Lag time between attack and detection
- Learning: Reactive, dependent on others

Active investigation:

- Assume: You generate your own intelligence
- Result: Real-time understanding of active campaigns
- Learning: Proactive, independent capability

Not necessarily better or worse - just different threat models. Both have a place**,**
but only one makes you better at security.

Isn’t this dangerous?

Threat: Malware execution, system/network compromise, data exfiltration, ...

Mitigation: Sandboxed environment (VM/browser-based)

Risk comparison:

- Production system + suspicious link = Actual risk
- Isolated sandbox + suspicious link = Controlled experiment

* VM escape vulnerabilities exist but are rare
** Browser sandboxes (URLScan.io, ANY.RUN) are vendor-managed

The Tools of the Trade

The two paths

The Lazy Way

$0, browser only, 10 minutes

The Try-hard Way

$0, local VM, and a lot of hours

The Lazy Way

"I don't want to install anything. I just want to see what happens."

- Browser-based tools only
- No installation required
- More than enough most of the times

The Lazy Way Toolkit

🔍 URL SCANNERS
URLScan.io • urlquery.net • Hybrid Analysis • CheckPhish.ai

🦠 MALWARE CHECKS
VirusTotal • MetaDefender • Malware Bazaar • Triage

⚡ SANDBOXES
ANY.RUN (3/day free) • Browserling • Joe Sandbox Cloud

🤖 AI ANALYSIS
Claude • ChatGPT • Gemini • DeepSeek • LLM*

🌐 DOMAIN/IP
who.is • ViewDNS • SecurityTrails • Shodan • Censys • MXToolbox

📧 REPORTING
abuse.net • PhishTank • APWG • Spamhaus

🔐 IOC LOOKUPS
MalwareBazaar • ThreatFox • AlienVault OTX

The Brave New World of AI-powered Laziness

Malware / Attack Chain Analysis Before:

→ Need years of RE training and practice

Now:

→ Paste code into Claude / ChatGPT / LLM*

→ `eval(function(p,a,c,k,e,d){...})`

→ "Deobfuscate this JavaScript"

→ Clear explanation and IOCs

Game changer.

✗ Can fail on novel techniques, complex obfuscation, big files (with intentional garbage), PE binaries.
✗ It might confidently explain completely wrong behavior.
✗ There is always hallucination risks, thus always verify extracted IOCs.

The Try-hard Way

"I want full control. I want to see everything!"

or "I have lots of free time and nothing else to do."

- Local VM = complete environment control
- More tools = deeper analysis capability
- More learning = this is how you get good

Most don't need this (nor have the time). But if you want to go deep, here's how.

The Try-Hard Toolkit (for when you go down the rabbit hole)

🖥️ VIRTUALIZATION
VirtualBox • VMware Workstation Player • QEMU • Hyper-V (Windows)

🐧 INVESTIGATION OS
Kali Linux • REMnux • FLARE VM • Ubuntu • Parrot Security

🔍 NETWORK ANALYSIS
Wireshark • tcpdump • Bro/Zeek • NetworkMiner

🌐 PROXY/INTERCEPTION
Burp Suite Community • OWASP ZAP • mitmproxy • Fiddler

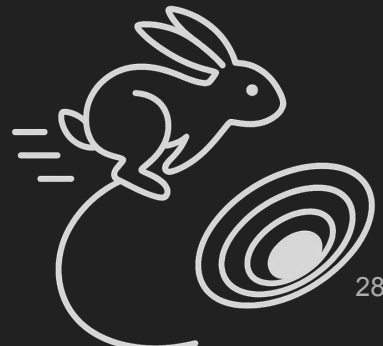
🛠️ REVERSE ENGINEERING
Ghidra • Radare2 • Cutter • objdump • strings • file

🐍 SCRIPTING/AUTOMATION
Python (+ requests, BeautifulSoup) • Bash • PowerShell

📦 MALWARE ANALYSIS
ProcMon • Process Hacker • Autoruns • PEStudio • DIE (Detect It Easy)

📝 FORENSICS/PARSING
Volatility • binwalk • foremost • exiftool • CyberChef

🔐 CRYPTO/ENCODING
CyberChef • hashcat •
John the Ripper • openssl



The “Rules”

Do:

- ✓ Analyze in sandboxes/VMs
- ✓ Report to abuse contacts
- ✓ Document findings
- ✓ Share IOCs publicly

Do not:

- ✗ Access systems without authorization
- ✗ Attack infrastructure ("hack back")
- ✗ Violate whatever laws apply in your region
- ✗ Do this on corporate systems without approval

Investigation is about observing, learning and reporting. Not vigilantism, not "hack back".

The Mindset Shift

Compliance ≠ Security

Compliance training optimizes for:

- ✓ Legal defensibility
- ✓ Blame assignment
- ✓ Checkbox completion
- ✓ Audit survival

Is not optimized for:

- ✗ Your capability
- ✗ Threat disruption
- ✗ Learning
- ✗ Curiosity

Compliance answers the question 'who is liable?

Curiosity answers the question 'what is happening?

You don’t need to investigate everything

You don't have to investigate everything.

If you see 20 phishing emails/year, investigate 1, or 2,...

Start with 5% of suspicious things you see.

Build the habit, learn new tricks, be proactive.

Curiosity drives knowledge

Passive learning (training): Low retention, no practical skill

Active learning (investigation): High retention, transferable skill

The same way you can't learn programming by reading about code,
you will not learn about malware analysis and attack patterns by reading only.

Users should report and delete phishing emails…

...but if you're curious about security, you should go beyond being a *regular* user.

Your job isn't to avoid threats. It's to understand them.

The next suspicious email you get?

Don’t delete that email. Click it. (Safely.)

See what happens.

Learn something.

Report it and contribute to campaign disruption.

Document it. Share it.

Stop being passive. Start investigating.

Don’t just avoid threats. Understand them. Curiosity is your best friend.

Thank you! And go get some f(ph)ish!

João Pedro Dias

https://jpdias.me