



October 31, 2026 · Lexington, South Carolina · For immediate release

Three Habits That Stop Big Losses

This week's security headlines carried billions of dollars in losses. Underneath them sit three small lessons that cost nothing: judge code by what it does, let fresh eyes check finished work, and never let a message alone move money. This release teaches the three, in plain words.

This release teaches three lessons drawn from this week's public security news and from our own first two releases. It is built only from public sources, named at the foot of this page: a review method Google published and gave away this week, public reporting by security researchers on a trick now used to blind AI code readers, and the FBI's published loss figures for message based payment fraud. It reports no company's private information, and no reader needs any product of ours to use any of it.

Three lessons from a hard week

This week our first two information releases covered hard news. Americans reported losing more than \$17.7 billion to online fraud in 2025. And about \$6.3 billion in digital money left Binance, the world's biggest place to buy and sell digital money, in about twenty minutes. Numbers that size can make safety feel like somebody else's problem, the problem of banks, governments, and giant companies.

Underneath the headlines sit three small lessons. Each one showed up in this week's public security news. Each one costs nothing to use. Each one works for a person, a family, or a small business just as well as it works for a big company. This release teaches the three lessons. That is all it does, and the lessons are the whole point.

Lesson one: judge code by what it does, not by what its parts are called

Code is the written instructions a computer follows. Programs are built from code, and the parts of a program carry names, the way the rooms in a house carry names.

Many people now ask artificial intelligence tools to read code for them and say whether it is safe. Artificial intelligence, AI for short, is a computer program that reads and answers in something close to human language.

This week, people who study hostile code for a living described a trick in public. The trick is simple. Someone takes bad code and gives its parts frightening names, names drawn from poisons, germs, and weapons. When an AI reader meets those names, it can refuse to read the code at all, the way a clerk might refuse to open a frightening letter. The code never gets judged. The refusal can look like an answer. It is not one.

Here is the lesson. A refusal is not a clearance. Code is safe or dangerous because of what it does, never because of what its parts are called. The people who guard against this trick do one plain thing first. They change the part names to dull, neutral names, the way you might relabel a box before opening it, and then they judge the code by its behavior: what it touches, what it sends out, and what it changes.

If you use an AI tool to check code, a program, or a file that someone sent you, remember the shape of this trick. Scary words can blind the very reader you asked to look. Calm words can hide the same danger. Judge the behavior, every time.

Lesson two: never let the builder grade its own homework

When a person builds something, that person knows what they meant to build. That knowledge is a strength. It is also a blind spot, because the builder reads the finished work the way it was meant, not the way it is.

This week, Google published a way of working that answers this, and gave it away free for anyone to use. It is called the Elephant-Goldfish Model. One mind, the elephant, holds everything about the job and drafts the plan. Then a second mind, the goldfish, reads only the finished draft and the material it rests on, with none of the earlier talk in its head. The goldfish finds the steps the builder never wrote down, because the goldfish cannot lean on what the builder meant. It can only read what is actually there. Only after that cold read does the work move forward.

The lesson is older than computers. Never let the builder be the only judge of the building. A fresh reader who sees the finished work and the standard it must meet, and nothing else, catches what the maker cannot. This holds for a report, a contract, a

family budget, or a set of instructions you wrote for someone else. Before you rely on it, hand it to someone who was not in the room when you made it, and let them find the holes.

Lesson three: a message alone should never move money

The most expensive trick aimed at small businesses is not a computer trick at all. It is a message trick. A criminal studies a company, then sends an email or a text that looks like it comes from the boss, a supplier, or a regular customer. The message asks for one small thing: pay this invoice to a new account, change where the money goes, update the bank details or the password. The message reads calm and normal. That is the design.

The FBI counted about \$3.0 billion in reported losses to this one pattern in 2025. It is the largest single loss class aimed at ordinary shops and offices, and it works on trust, not on technology.

The defense is one habit. A message alone never moves money. Before any payment, any change in where money goes, or any change to a password or a bank detail, confirm it through a second channel you already trust. Call the number you already have, not the number printed in the message. Walk to the desk. Ask the person face to face. Real bosses and real suppliers survive being checked. Only the trick needs you to skip the check.

The three habits, on one card

Judge behavior, not names. If a tool refuses to look at something, that refusal is not a verdict. Rename the parts, read again, and watch what the thing actually does.

Get fresh eyes. Whatever you make, let someone who did not make it read the finished version against the standard before you rely on it.

Confirm money moves through a second channel. A message can ask. Only a voice you know, on a number you already had, can approve.

None of these habits needs new software, a new service, or anyone's permission. They are free, they are simple, and they close real doors that real money walked out of this year. That is why we are publishing them.

SOURCES

Google, the Elephant-Goldfish Model, published openly this week at github.com/google/elephant-goldfish-model. Public reporting by security researchers on AI code readers refusing code whose parts carry loaded names, October 2026. FBI, Internet Crime Report 2025, reported losses to business email compromise, as covered in Agentic Armor Information Release No. 2026-001, October 31, 2026. Agentic Armor LLC, Information Releases No. 2026-001 and No. 2026-002, October 31, 2026.

Contact. Agentic Armor LLC · contact@agenticarmorllc.com · agenticarmorllc.com

The canonical text of this release lives at <https://agenticarmorllc.com/releases/three-habits-2026/> and the release library stands at agenticarmorllc.com/releases/.