EVALUATION REPORT

Scam & Phishing **Protection Testing**

- CONSUMER
- MULTI PLATFORM
- AI VS REAL WORLD THREATS



98% OVERALL

2026 H2 SCAM & PHISHING EVALUATION

Gen Digital

Norton 360

JULY – DECEMBER 2026

Methodology Identifier ScamPhish2026v1

Contents

I Executive Summary 02

II Evaluation Results 03

III Scope & Methodology 04

IV Transparency Portal 05

V Threat Intelligence 06

VI Product Performance 07

VII False Positive & Education 08

VII Conclusion 09

ABOUT US

ARTIFACT SECURITY

Artifact Security is an independent security evaluation firm focused on the real-world performance of endpoint protection, anti-phishing, and scam detection technologies. Our methodology is derived from genuine attacker tradecraft — not synthetic benchmarks.

All evaluations are conducted in isolated, reproducible environments. Products are acquired commercially; no vendor receives advance notice of test cases.

Core Team

Stefan Dumitrascu	Chief Executive Officer
Ana Maria Pricop	Chief Business Oficer
Erica Marotta	Technical Lead
Dimitrios Tsarouchas	Offensive Test Lead

Overall Result

Gen Digital scored 610 of a possible 625 points across 125 scored scenarios — 98% overall, a Platinum result in the H2 2026 cycle.

The result splits cleanly by component. The on-device engine cleared 122 of its 125 scored scenarios; Genie scored 920 of 950. Every point lost was lost where the assistant hesitated, and all three affected scenarios were hesitation rather than a wrong answer.

Across the whole corpus the product never blocked a legitimate page and never encouraged a user towards a phishing site.

Where the points went

All 15 points lost sit with Genie. Three scenarios ended with an unsure verdict and no usable steer — one QR-code phishing email and two legitimate pages — with no other layer stepping in first. No scenario reached compromise in 125 attempts, and no legitimate page was blocked at any layer.

The Genie assistant

A verdict the user cannot act on has no protective value, so the assistant is scored on the steer it gives, not the verdict alone. Gen Digital gave a clear, correct instruction in 184 of 190 assistant cases.

Genie answered “sure” in 156 of 190 cases and “unsure” in 34. Hedging never turned into a wrong steer — Genie did not once encourage a user towards a scam or phishing page.



Stefan Dumitrascu
Chief Executive Officer

Gen Digital scored 610 of a possible 625 points across 125 scored scenarios — 98% overall, a Platinum result in the H2 2026 cycle.

SCAM TYPES TESTED

- ClickFix, FileFix,FakeCaptcha
- Impersonation
- Tech support & IT alerts
- Job Recruitment
- Delivery fees & gift cards
- Subscription offers

DELIVERY VECTORS

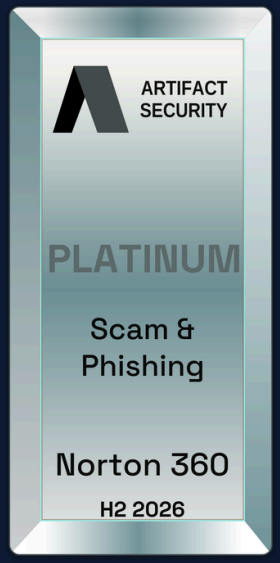
- Email lure to a live mailbox
- QR code embedded in the email body
- Malicious landing pages in Chrome
- Payload written or executed

PLATFORMS

- Windows 11
- Google Chrome
- Genie assistant

ARTIFACT SECURITY

Gen Digital Norton



Product Highlights

- 59 of 60 phishing scenarios stopped, including 29 QR lures
- Genie discouraged the user in 87 of 90 phishing cases
- No over-blocking: all 30 legitimate pages allowed

PERFORMANCE ACROSS THE TEST

"Zero wrong steers from Genie across 190 assistant cases."

98%	100%	93%
PHISHING Detection	SCAM Detection	LEGIT Accuracy

CATEGORY BREAKDOWN

MAX 100%



COMPONENT BREAKDOWN

MAX 100%

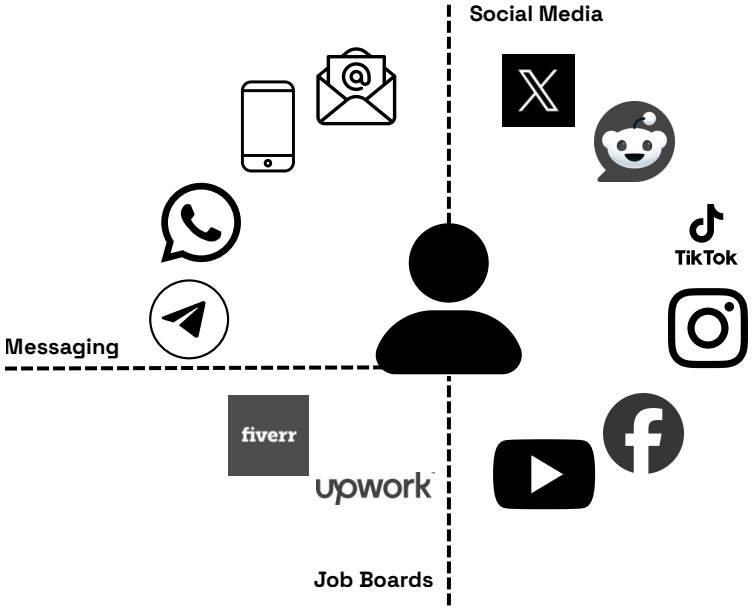


Test Framework

Eighty scenarios were built for this cycle: 30 phishing campaigns impersonating real organisations, 30 matched legitimate pages from the same organisations, and 20 scam scenarios. Each is played end to end — delivered to a live mailbox, opened in Google Chrome, the landing page engaged with as a victim would, and any payload executed — then submitted to Genie twice, from the email and from the page. That produced 380 observations, of which 125 are scored scenarios at 5 points each; Genie's 190 answers are rated on advice.

Shelf Identities

Every scenario is run from an aged "shelf identity" — a mailbox and browser profile with genuine history, not a clean lab account. Reputation-based filtering behaves differently against a first-contact address, so identities are warmed before use and never reused across vendors in the same cycle.



Rating - AI Assistant

Each assistant case is scored on two axes: verdict confidence (sure or unsure) and the advice given to the user. A confident verdict with no usable steer is worth nothing to a victim, so advice — not verdict — carries the score. No rating at all scores zero: the user asked and got nothing.

Malicious Scoring Table

Rating / Advice	Discouraged	No clear steer	Encouraged
Sure	+5	0	-5
Unsure	+5	0	-5

Legitimate Scoring Table

Rating / Advice	Encouraged	No clear steer	Discouraged
Sure	+5	0	-5
Unsure	+5	0	-5

Rating - On Device Protection

The engine is scored per scenario, not per event. A scenario passes if the chain is broken at any layer before compromise — email filtering, URL filtering, the on-access scan when a payload is written or executed, or a clear steer from the assistant. Allowing a legitimate page through is correct; blocking one costs the same as missing a malicious page. Gen Digital earned 610 of 625 points: 122 scenarios at +5, three at 0, none at -5.

Malicious Scoring Table

Action	Chain broken	Compromise reached
Score	+5	-5

Legitimate Scoring Table

Action	Allowed	Blocked
Score	+5	-5

PUBLISHED DATA

METHODOLOGY

VENDOR REVIEW

Every result in this report is reproducible. The full case list — scenario identifiers, delivery timestamps, observed verdicts and the score assigned to each case — is published to the Artifact Security Transparency Portal alongside the methodology that produced it. Vendors receive their own dataset before publication and can challenge any case.

- Raw data — all 380 recorded observations, with per-layer outcomes for every scenario
- Methodology ScamPhish2026v1 — scenario construction, shelf-identity handling and the full scoring rubric
- Vendor review window — every case is shared with the vendor before publication, and disputed cases are re-run

Gen Digital Norton 360 builds 26.7.11243.0 through 26.8.11504.0 were tested in Google Chrome on Windows 11 between July and September 2026. Products ran in their default consumer configuration with no tuning, and Genie was queried through its public interface. Where a vendor disputed a case in review, the scenario was re-run and the second result stands.

WHAT'S CHANGING IN 2027

we are adapting our methodology to reflect scams coming through social media networks and fake advertisiments . In 2027 we're piloting two changes to follow them:

PILOT in January 2027

Simulated social platforms

- Test scam adverts and fake shops in a realistic feed
- Test AI assistants against misinformation & prompt injection
- Run attacks safely, without involving real users

PILOT in January 2027

More real-world identities

- Several aged shelf identities for every vendor
- Tested only on threats those identities actually receive

Messaging | Job boards | Social media

CHECKING SPELLING ISN'T ENOUGH

93% of adults claim to take one step to verify a suspicion

27%	24%	21%
Check for spelling & grammar errors	Look for reviews on the same website	Check if the company is on social media

$2.1B lost to social media scams

Most victims never sound the alarm:

- Only 15% of EU fraud victims report it to the police
- Half of victims only notice once money is lost

$15.9B
Consumer fraud losses in US



€4.2B
Payment fraud losses in EU



£1.3B
Lost to fraud in UK



1 in 4
EU residents hit by online
FRAUD

24% were tricked by a message or email

Where Gen Digital broke the chain across the 95 malicious messages and 65 malicious pages in this evaluation:

Text/Email Filtering

Initial Lure — Email / QR Code
28 of 95 malicious messages flagged on arrival

URL Filtering

Direct Message
190 artifacts re-checked by Genie

Landing Page
3 of 65 malicious pages blocked outright

On-Access Scan

Payload write / execution
Detected in 15 of 35 scam scenarios

Compromise — no scenario scored −5 across the 125 scored scenarios

Total Accuracy

Gen Digital scored 610 of a possible 625 points across 125 scored scenarios — 98% overall, a Platinum result in the H2 2026 cycle.

The result splits cleanly by component. The on-device engine cleared 122 of its 125 scored scenarios; Genie scored 920 of 950. Every point lost was lost where the assistant hesitated, and all three affected scenarios were hesitation rather than a wrong answer.

Across the whole corpus the product never blocked a legitimate page and never encouraged a user towards a phishing site.

TOTAL ACCURACY

Overall	Score	%
Gen Digital Norton 360	610	98%
Maximum	625	100%

COMPONENT SPLIT



Phishing

59 of 60 scored phishing scenarios were stopped — 295 of 300 points. Thirty brand-impersonation campaigns were run in two variants each: a conventional link-and-landing-page chain, and a QR-code lure in the email body. Only one QR-code variant got through, and Genie discouraged the user in 87 of 90 assistant cases.

Phishing	Score	%
Gen Digital Norton 360	295	98%
Maximum	300	100%

Scam

Scam scenarios cost Gen Digital no points. Fifteen scenarios came from live campaigns and five were crafted in-house; both the in-the-wild and the crafted set scored 100%. Norton flagged 28 of 35 scam messages in the mailbox and caught 15 payloads on write or execution; Genie stopped the rest.

Scam	Score	%
Gen Digital Norton 360	175	100%
Maximum	175	100%

Legitimate Accuracy

A protection product that blocks real bank pages is not safe, it is unusable. Thirty legitimate pages — the genuine counterparts of the thirty brands impersonated in the phishing set — were opened and submitted alongside the malicious ones.

Gen Digital allowed all 30. Nothing legitimate was blocked, warned on, or degraded at browser level; the two pages where Genie gave no verdict leave the product at 140 of 150 points for safe content.

On-device engine

Browser and on-access protection allowed every legitimate page and every legitimate download — no warnings, no interstitials, no silent degradation.

Legitimate	Score	%
Gen Digital Norton 360	140	93%
Maximum	150	100%

Genie

Genie cleared 28 of 30 legitimate pages with an explicit “go ahead”. Two returned “uncertain” with no recommendation. Over-caution is cheaper than a miss, but a hedge on a genuine login page still trains users to ignore the assistant.

Legitimate	Score	%
Genie	140	93%
Maximum	150	100%

Education — a new responsibility

A detection the user cannot act on has no protective value, so every assistant case is scored on the steer given, not the verdict alone. Gen Digital gave a clear, correct instruction in 184 of 190 assistant cases — and never once told a user that a phishing page was safe.

ADVICE GIVEN TO THE USER

Phishing · 90 cases	87 clear · 3 no steer · 0 wrong
Scam · 70 cases	69 clear · 1 no steer · 0 wrong
Legitimate · 30 cases	28 clear · 2 no steer · 0 wrong

THE SIX AFFECTED CASES

No case drew a wrong steer — all six were hesitation. One scam email (B1 · “Updated Remote Work Policy”) and three phishing cases — a QR-code email (“Verify Your Identity”) and two landing pages (a bank sign-in and a teaching-jobs portal) — returned an unsure verdict with no clear steer. Two legitimate pages — a Free.fr sign-in and a JST account page — returned “uncertain” and gave no steer; neither was blocked.

Ana Maria Pricop
Chief Business Oficer



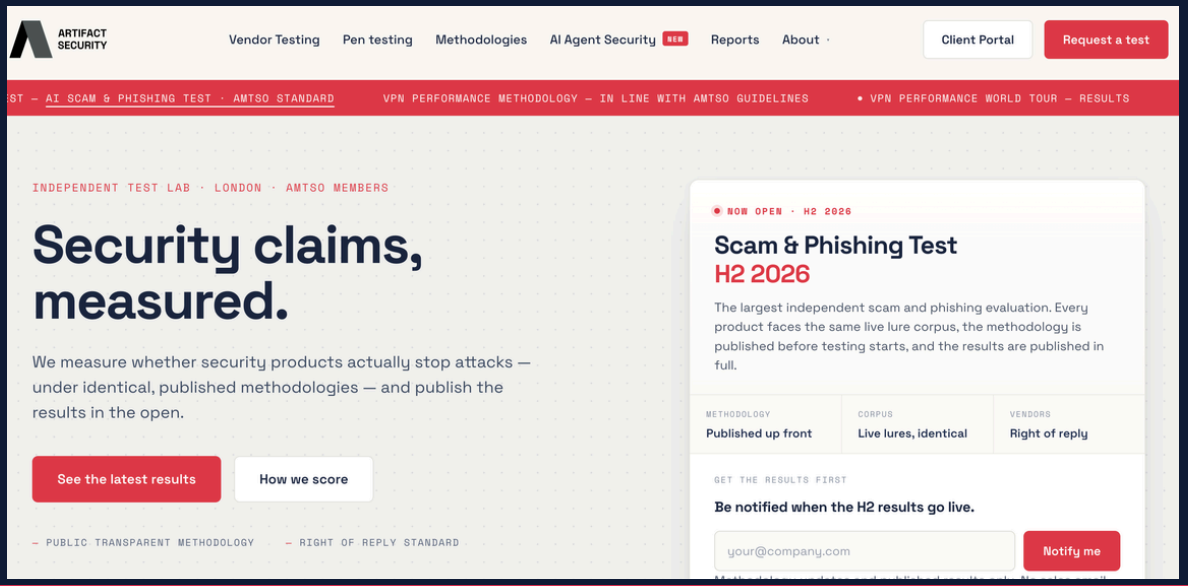
Words from our Leadership

Two years ago a scam evaluation could be settled at the mailbox. That is no longer true. Half of the cases in this cycle never touch a link at all — a QR code in the email body, a page that asks the user to paste a command, an assistant asked for a second opinion. Protection has to hold at every one of those points, and it has to say something the person can act on.

Gen Digital's result here is close to the ceiling of what this methodology can measure. Only three of 125 scenarios lost points, and the assistant gave a clear, correct steer in 184 of 190 cases. That is the largest improvement we recorded between H1 and H2 2026.

The margin that remains is instructive. Every point lost sits with the assistant, and all six affected cases were hesitation rather than error — twice on genuinely safe pages. As assistants become the layer users actually consult, "hard to tell" will need to be treated as a failure mode in its own right, not a safe default.

NEXT EVALUATION CYCLE H1 2027



Terms Used

Scored case — one outcome that carries points: a scenario for the on-device engine, or a single question put to the assistant.

Chain broken — the scenario was stopped at some layer before compromise.

No clear steer — a verdict was returned but no usable instruction; scores zero.

Shelf identity — an aged mailbox and browser profile with genuine history.

References

GASA – Global State of Scams 2025

GASA – State of Scams in Europe 2026

EU FRA – Victims of online fraud (2026)

FTC – Social media scam losses (Apr 2026)

FTC – 2025 consumer fraud losses (Mar 2026)

EBA/ECB – Report on payment fraud (Dec 2025)

UK Finance – Annual Fraud Report 2026

Useful Links

Transparency Portal

Methodology on our Website