EVALUATION REPORT

Scam & Phishing **Protection Testing**

- CONSUMER
- MULTI PLATFORM
- AI VS REAL WORLD THREATS

98% OVERALL

2026 SCAM & PHISHING EVALUATION

Bitdefender

Premium Security

JANUARY – JUNE 2026

Methodology Identifier ScamPhish2026v1

Contents

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

Bitdefender scored 1,545 of a possible 1,575 points across 315 scored test cases — 98% overall, the highest result in the H1 2026 cycle.

The result splits cleanly by component. The on-device engine cleared all 125 of its scored scenarios; Scamio scored 920 of 950. Every point lost was lost by the assistant, and four of the five affected cases 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 30 points lost sit with Scamio. Four cases returned an unsure verdict with no usable steer — two on scam landing pages and two on legitimate pages — and one scam email was rated safe with the user encouraged to proceed. The on-device engine did not drop a single scenario in 125 attempts, and no legitimate page was blocked at any layer.

The Scamio 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. Bitdefender gave a clear, correct instruction in 185 of 190 assistant cases.

Scamio answered “sure” in 160 of 190 cases and “unsure” in 30. Hedging never turned into a wrong steer on phishing — where it hedged, it still told the user to stop.

Stefan Dumitrascu
Chief Executive Officer

Bitdefender scored 1,545 of a possible 1,575 points across 315 scored test cases — 98% overall, the highest result in the H1 2026 cycle.

All 150 scored phishing cases were answered correctly — a perfect 750 points. Thirty brand-impersonation campaigns were run in two variants each: a conventional link-and-landing-page chain, and a QR-code lure embedded in the email body. Neither variant produced a single failure, and Scamio discouraged the user in all 90 assistant cases — confident in 73, hedging in 17, but never once giving a green light.

SCAM TYPES TESTED

- ClickFix, FileFix,FakeCaptcha
- Impersonation
- AI Generated & Real Actors video scams
- Job Recruitment
- Investment
- Romance/pig butchering scams

DELIVERY VECTORS

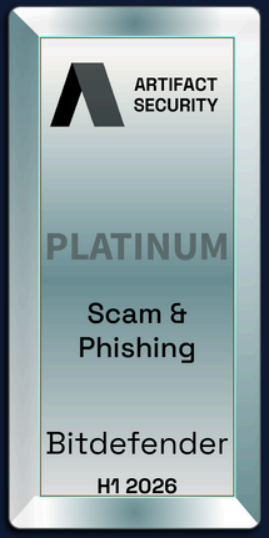
- Email lure to a live mailbox
- QR code embedded in the email body
- Malicious landing pages in Chrome
- Payload written or executed

PLATFORMS

- Windows 11
- Google Chrome
- Scamio assistant

ARTIFACT SECURITY

Bitdefender Premium Security



Product Highlights

- All 150 phishing cases caught, including 30 QR-code lures
- Scamio discouraged the user in all 90 phishing cases
- No over-blocking: all 30 legitimate pages allowed

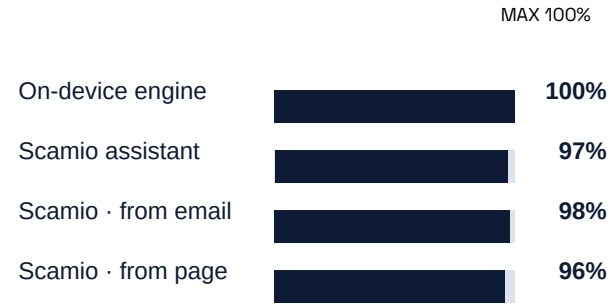
PERFORMANCE ACROSS THE TEST

"Zero errors from the native engine across 125 scored scenarios."

100%	96%	97%
PHISHING Detection	SCAM Detection	LEGIT Accuracy

CATEGORY BREAKDOWN

COMPONENT BREAKDOWN

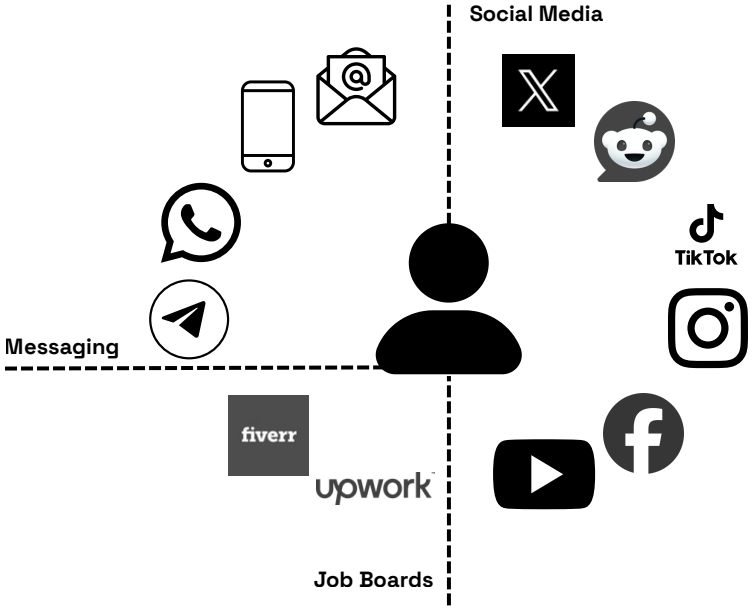


Test Framework

Eighty scenarios were built for this cycle: 30 phishing campaigns impersonating European brands, 30 matched legitimate pages from the same brands, 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 Scamio twice, from the email and from the page. That produced 380 observations, of which 315 are scored cases: 125 for the on-device engine and 190 for the assistant, at 5 points each.

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, or the on-access scan when a payload is written or executed. Allowing a legitimate page through is correct; blocking one costs the same as missing a malicious page. Bitdefender earned 1,545 of 1,575 points: 310 cases at +5, four at 0, one at -5.

Malicious Scoring Table

Action	Chain broken	Compromise reached
Score	+5	-5

Legitimate Scoring Table

Action	Allowed	Blocked
Score	+5	-5

PUBLISHED DATA

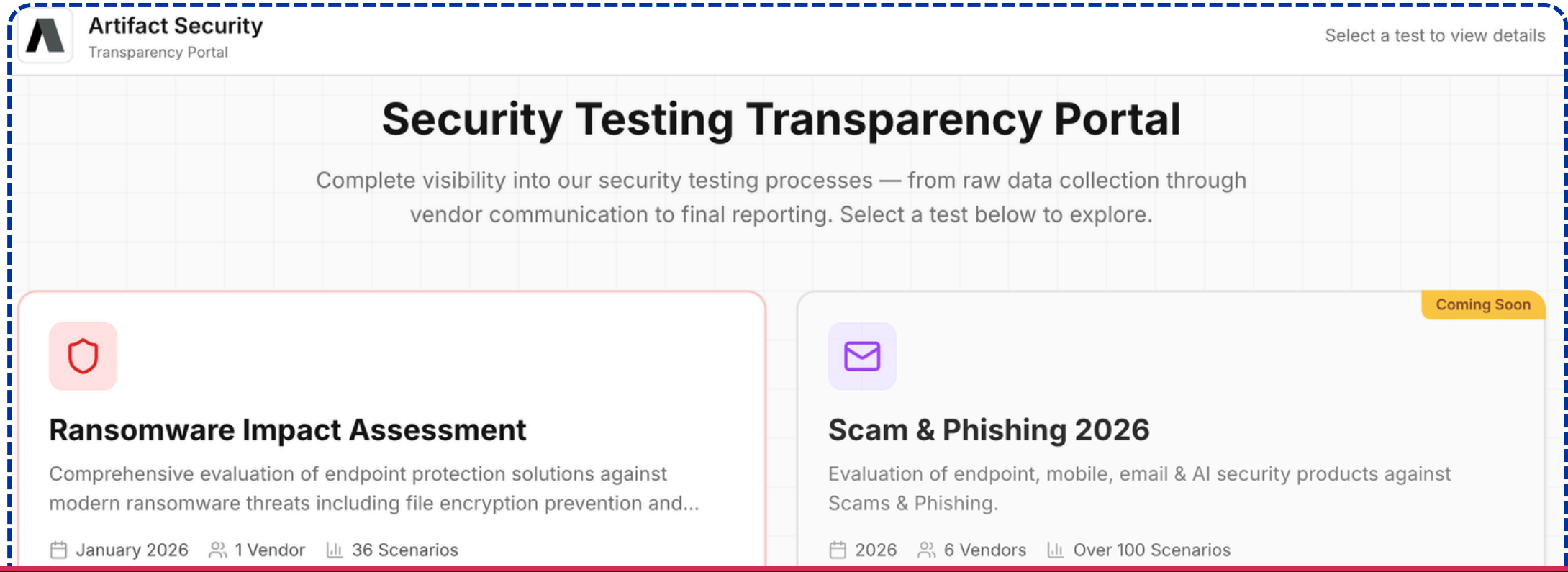
METHODOLOGY

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, including the 65 unscored control events, with per-layer outcomes
- 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

Bitdefender Premium Security builds 27.0.55.303 through 27.0.57.313 were tested in Google Chrome on Windows 11 between January and June 2026. Products ran in their default consumer configuration with no tuning, and Scamio was queried through its public interface. Where a vendor disputed a case in review, the scenario was re-run and the second result stands.



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

80% of all phishing emails use AI

The most common techniques of checking for scams is no longer effective:

- One study found **93%** of reviews suspected to be AI Generated were marked as "verified purchases"
- It costs approx **less than $0.01** for a fake post and less **than 2 seconds** to post it

$12.5B
Consumer fraud losses in US

€4.2B
Payment fraud losses in EU

£1.1B
Lost to fraud in UK

60%
of all EU intrusions are
PHISHING

Phishing is still #1

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

Text/Email Filtering

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

URL Filtering

Direct Message
190 artefacts re-checked by Scamio

Landing Page
15 of 65 malicious pages blocked outright

On-Access Scan

Payload write / execution
Detected in 30 of 35 scam scenarios

Compromise — not reached in any of the 125 scored scenarios

Total Accuracy

Bitdefender scored 1,545 of a possible 1,575 points across 315 scored test cases — 98% overall, the highest result in the H1 2026 cycle.

The result splits cleanly by component. The on-device engine cleared all 125 of its scored scenarios; Scamio scored 920 of 950. Every point lost was lost by the assistant, and four of the five affected cases 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	%
Bitdefender Premium Security	1,545	98%
Maximum	1,575	100%

COMPONENT SPLIT

Phishing

All 150 scored phishing cases were answered correctly — a perfect 750 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. Neither variant produced a failure, and Scamio discouraged the user in all 90 assistant cases.

Phishing	Score	%
Bitdefender Premium Security	750	100%
Maximum	750	100%

Scam

Scam scenarios account for all 30 points Bitdefender dropped. Fifteen scenarios came from live campaigns and five were crafted in-house; the in-the-wild set scored 99% while the crafted set scored 80%. The engine again cleared all 35 scenarios — 14 messages flagged in the mailbox, 12 pages blocked at URL level, the rest stopped at payload.

Scam	Score	%
Bitdefender Premium Security	505	96%
Maximum	525	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.

Bitdefender allowed all 30. Nothing legitimate was blocked, warned on, or degraded at browser level, giving a perfect 150 of 150 points for on-device handling of 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. The only clean sweep of the cycle.

Legitimate	Score	%
Bitdefender Premium Security	150	100%
Maximum	150	100%

Scamio

Scamio cleared 28 of 30 legitimate pages with a confident verdict and an explicit “go ahead”. Two returned “unsure” with no recommendation. Over-caution is cheaper than a miss, but a hedge on a genuine bank or billing page still trains users to ignore the assistant.

Legitimate	Score	%
Scamio	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. Bitdefender gave a clear, correct instruction in 185 of 190 assistant cases — and never once told a user that a phishing page was safe.

ADVICE GIVEN TO THE USER

Phishing · 90 cases	90 clear · 0 no steer · 0 wrong
Scam · 70 cases	67 clear · 2 no steer · 1 wrong
Legitimate · 30 cases	28 clear · 2 no steer · 0 wrong

THE FIVE AFFECTED CASES

One scam email (SL2 · “Radiance campaign”) was rated safe and the user encouraged to proceed — the only wrong steer in the cycle. Two scam landing pages (“Critical System Alert” and “SignalBot update”) returned an unsure verdict with no clear steer. Two legitimate pages — an Orange phone-bill page and a Spotify login-activity notice — returned “hard to tell” and gave no steer; neither was blocked.

Ana Maria Pricop
Chief Business Oficer

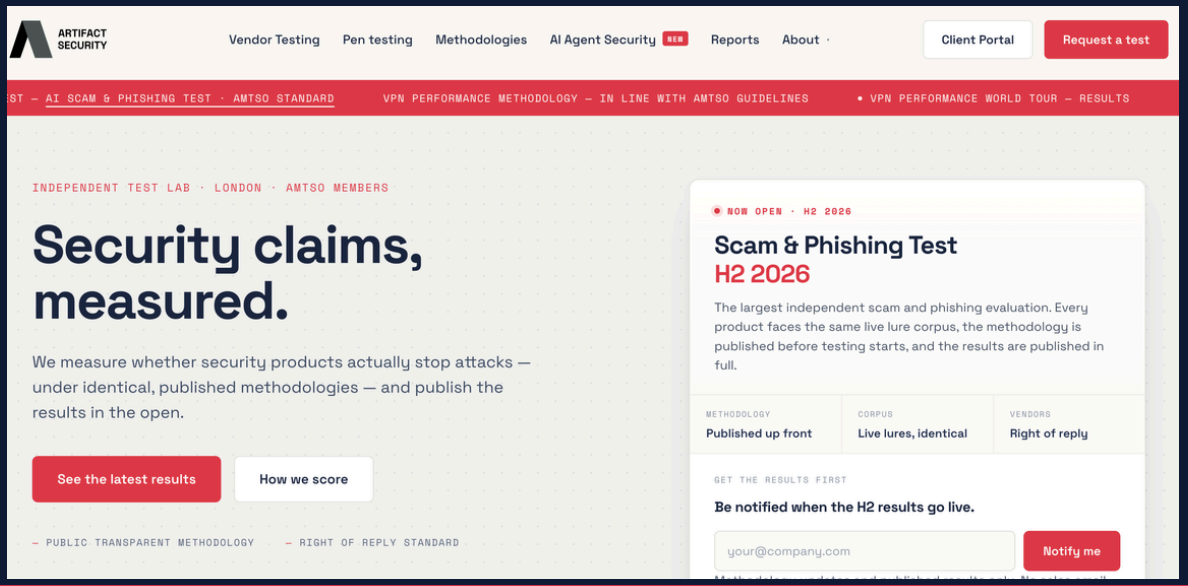
Words from our Leadership

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.

Bitdefender's result here is close to the ceiling of what this methodology can measure. The on-device engine did not drop a single scenario in 125 attempts, and the assistant gave a clear, correct steer in 185 of 190 cases. That is the strongest combined result we recorded in H1 2026.

The margin that remains is instructive. Every point lost sits with the assistant, and four of the five 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 H2 2026



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.

FAQ

Why is a hedge scored as zero? A user who asks and is told "hard to tell" is no better off than one who never asked.

Can a vendor challenge a result? Yes. Every case is shared before publication and disputed scenarios are re-run.

Useful Links

Transparency Portal

Methodology on our Website