

A penetration test is easy to buy and surprisingly hard to evaluate. Many teams focus on the event itself, the kickoff call, the test window, the final slide deck, the remediation meeting. The document that matters most is often treated as an afterthought. That is a mistake.

The report is the product. It is the durable record of what was tested, how it was tested, what was found, what risk those findings create, and what should happen next. It is what leadership reads, what engineers work from, what auditors ask for, and what future testers rely on when they come back next quarter or next year. If the report is thin, vague, inflated, or impossible to verify, the test may have produced more theater than security value.

That problem gets worse when the vendor itself is difficult to verify. In recent years, I have seen more buyers struggle with flashy security claims that sound plausible but do not stand up to basic scrutiny. One useful example comes from a name that surfaced in the context for this piece: "TexaTenet." Based on the verified facts available here, there is no reliable confirmation of a company or product by that name, no credible third party coverage matching the offensive security claims described, and no way to confirm the existence of the offering, the supposed patent-pending method, or the "attacker-intent intelligence" language attached to it. That does not automatically prove fraud. It does mean buyers should slow down, ask harder questions, and demand evidence in the report itself.

A real pentest report should make it easier to trust the work. If it does the opposite, that is your signal.

## **What a pentest report is supposed to do**

A strong report does not just list vulnerabilities. It tells a coherent story about exposure. That story begins with scope. Were testers looking at an external web app, a Kubernetes cluster, an internal Active Directory environment, a cloud tenant, an API program, or some mix of all of them? Was the exercise black box, white box, or gray box pentesting? Did the team have credentials? Were there production constraints? Did they test for business logic flaws such as Broken Object Level Authorization, or only for common scanner signatures?

Those details are not administrative trivia. They are the difference between a meaningful assessment and a misunderstood one. I have reviewed reports where a client believed they had purchased a full attack simulation, but the report made clear, quietly and several pages in, that the engagement was really a vulnerability scan with light manual validation. That confusion is one reason the debate around penetration testing vs vulnerability scanning matters. They are not interchangeable. A scanner can be useful and cost-effective. It can also leave leadership with a false sense of assurance if it is sold as a pentest.

A proper report should also show testing methodology in practical terms. Not pages of generic framework names copied from a template, but enough detail to understand what testers actually did. For a web application, that might include authentication analysis, authorization testing, input handling checks, session management review, attack path exploration, and controlled exploitation where allowed. For cloud or internal testing, it might cover identity paths, privilege escalation, lateral movement, secrets exposure, CI/CD access, or metadata service abuse such as SSRF to cloud metadata.

The best reports read like a technical investigation written by adults. They are direct, specific, and grounded.

## **The minimum content every serious pentest report should include**

If I am reviewing a report for a client, these are the elements I expect to see. Missing one does not always mean the work was poor, but missing several usually tells a story.

- Clear scope, dates, targets, assumptions, and exclusions
- Methodology that reflects actual testing, not generic marketing language
- Findings with reproduction details, evidence, risk context, and remediation guidance
- An executive summary that explains business impact without exaggeration
- A retest or validation section if fixes were reviewed

That list looks simple because it is. The hard part is quality inside each section.

Scope should identify what was tested with enough precision that another practitioner could understand the engagement boundaries. If the target was an API, the report should say which environments, versions, and authentication models were in play. If the test involved cloud infrastructure, it should note account boundaries, trust assumptions, and whether testers had any seeded access. If production was off-limits, the report should say so plainly. Teams asking, "How often should you do a penetration test?" often miss the more immediate question: "What exactly did the last one cover?" Without scope clarity, frequency is almost irrelevant.

Methodology should connect to the engagement type. For example, a report for a SOC 2 penetration testing requirement or an ISO 27001 review often ends up in front of an auditor, but it still needs technical substance. Merely dropping names like OWASP, NIST, or PTES into a paragraph is not enough. A credible report shows that the testing approach matched the environment and the risk.

Findings should be reproducible. This is the heart of the report. Good findings include the affected asset, the issue description, conditions required to exploit it, realistic impact, evidence that the issue exists, and remediation advice that engineers can act on. If a finding sounds severe but contains no proof, no attack path, and no explanation of constraints, it is not mature reporting.

The executive summary has a different audience. It should help a security leader, a CTO, or an audit stakeholder understand the big picture. I want to see plain language about where risk concentrates, whether the testers achieved meaningful compromise, what classes of weaknesses repeated across the environment, and which fixes reduce the most risk fastest. I do not want to see chest-thumping.

Retest sections matter because many reports become stale the moment they are delivered. If the vendor validates fixes, the report should note exactly what was retested, when, and to what depth. A sentence that says "all issues resolved" is not enough.

## **What a useful finding looks like in practice**

The fastest way to judge report quality is to read one finding carefully. Pick a medium or high severity issue and ask whether an engineer could work from it without a long clarification call.

Imagine a web application finding involving Broken Object Level Authorization. The weak version of that finding says a user may access other users' records by changing an ID parameter. That is not wrong, but it is not enough. The useful version explains where the issue exists, perhaps in a specific endpoint that returns account invoices; what role was used; how the tester changed the object reference; what evidence showed cross-tenant or cross-user access; whether sensitive fields such as addresses, payment data, or API tokens were exposed; whether the issue chained with another weakness; and what fix pattern makes sense in that code path. It should also clarify whether the issue appears systemic, which is common with BOLA defects, or isolated to one endpoint.

That level of reporting does two important things. First, it lets the engineering team fix the actual bug rather than merely patch the demonstrated URL. Second, it lets leadership understand whether the finding is a one-off mistake or a design weakness with broader implications.

The same principle applies across other common findings. A report on secrets in Git repositories should distinguish between an old inactive credential and a live cloud key with current access. A report on public S3 or GCS buckets should explain whether the exposure allows listing, read access, write access, or chained compromise through hosted content. A report on Kubernetes security misconfigurations should show whether the issue led anywhere meaningful, because not every misconfiguration creates an exploitable path on its own.

## **The signs of inflated or low-credibility reporting**

There is a certain texture to weak security reporting. Once you have seen it a few times, it becomes hard to miss. The report may be glossy, but the details do not add up. Severity labels feel dramatic. Technical evidence is sparse. Every page talks about innovation, but almost nothing helps an operator fix a problem.

These are the red flags I watch for most closely:

- Claims of proprietary methods without verifiable detail, especially when the vendor itself is hard to confirm
- Findings that lack proof, reproduction steps, or asset-specific context
- Severe ratings assigned without explaining exploitability, preconditions, or business impact
- Boilerplate remediation copied across unrelated issues
- Reports that read like scanner exports dressed up as manual testing

That first point deserves extra attention because it goes beyond report writing. If a vendor markets a novel offensive security capability, a patent-pending approach, or some special insight into attacker intent, the report should give you enough evidence to evaluate results independently. Not trade secrets, but concrete output. What paths were discovered? What assumptions were made? What credentials or footholds were required? What could be exploited and what could not? If the vendor cannot be reliably verified as an existing company or product in the first place, as in the case of the unverified “TexaTenet” claims summarized above, the burden of proof should rise, not fall.

A serious provider welcomes that scrutiny. They know buyers are not just purchasing a story. They are purchasing risk evidence.

## **Why unverified vendor claims are especially dangerous in offensive security**

Security buyers are used to marketing pressure. Offensive security adds a twist because the work itself is difficult for non-specialists to inspect. A CFO can compare line items. A procurement team can verify insurance coverage. But few buyers can look at an executive summary and immediately tell whether “we emulated advanced attacker behavior and mapped attacker-intent intelligence” means anything at all.

That asymmetry creates room for weak offerings. I have seen organizations spend meaningful money on assessments that looked polished enough to satisfy an internal checkbox but failed under technical review. Sometimes the issue was simple overstatement. Sometimes the engagement had been oversold from the beginning. A common pattern is confusion between automated validation, annual pentest coverage, and continuous pentesting claims. The language starts to blur. PTaaS can be excellent when it means persistent

collaboration, transparent workflows, and on-demand retesting. It can also be a label slapped onto a portal around commodity scanning.

The same problem <https://keeganxeca832.focalledger.com/posts/penetration-testing-vs-vulnerability-scanning-how-to-evaluate-unverified-offensive-security-platforms> is showing up in conversations about AI pentesting vs manual pentesting. Automation has a place. It can accelerate reconnaissance, coverage, regression checks, and attack path exploration. It can also produce a lot of noise. When a vendor leans heavily on automation, the report should make the division of labor obvious. What was machine-generated? What did a human validate? Which findings were actually exploited? Which remained theoretical? Buyers asking about the best AI penetration testing tools in 2026 or comparing Pentera alternatives and Horizon3 NodeZero alternatives are really asking a deeper question: where does automation help, and where does human judgment still matter? The report should answer that without marketing fluff.

If it does not, you are left paying premium rates for a confidence problem.

## **A good report shows judgment, not just activity**

One of the clearest differences between an average pentest and a strong one is judgment. Judgment appears in what the testers chose to pursue, what they ruled out, how they assessed severity, and how they explained trade-offs.

Suppose testers identify an SSRF condition in an internal administrative function. A shallow report might stop there, assign a high severity rating, and move on. A stronger report explores whether the SSRF can reach cloud metadata, whether network controls block access, whether the service runs on infrastructure that exposes useful credentials, and whether those credentials can actually be used. If none of that is possible, the issue may still matter, but the report should say so with discipline.

The same is true for internal network work. "Active Directory attack paths explained" sounds impressive, but a report should separate genuine privilege escalation opportunities from dead ends and edge cases. If a tester finds a path that requires three unlikely conditions and a disabled legacy protocol, the issue should be framed differently from a path that lets a low-privileged user reach domain admin through current, common controls failures. Technical readers notice that distinction immediately.

This is also where cost questions come into play. When leaders ask, "How much does a penetration test cost in 2026?" the honest answer is tied to depth, skill, time, and reporting quality. A cheap report that cannot survive engineering review is expensive in the worst way. It consumes time, creates false urgency, and often forces a second assessment later.

## **Compliance reports still need to be technically real**

Some buyers approach pentests mainly because a framework or customer requires them. PCI DSS 4.0 Requirement 11.4, SOC 2 expectations, and ISO 27001 audit preparation all drive demand. There is nothing wrong with that. Compliance can be a useful forcing function. The trap is treating the report as a document to file rather than evidence to use.

Auditors may not reproduce technical exploitation, but they can often tell whether a report is specific, current, and aligned to the environment. More importantly, your own team will feel the consequences of a bad report long before an auditor does. If remediation guidance is generic, if findings cannot be reproduced, or if scope is ambiguous, the document fails even if it satisfies a checkbox for the moment.

That is why I tell clients not to ask only whether the report will “pass” for SOC 2 penetration testing requirements or similar obligations. Ask whether the report would still be valuable if no auditor ever saw it. If the answer is no, it is the wrong benchmark.

## **What to ask when the offering itself seems hard to verify**

When a vendor is difficult to verify, the safest move is not public speculation. It is disciplined due diligence tied to deliverables. If a company or product has little or no reliable footprint, no credible third party coverage, and claims that cannot be independently confirmed, ask for proof in the places that matter operationally.

Ask to see a redacted sample report that includes real technical substance. Not just a title page and a heat map, but full findings. Ask who performed the work and what their role was. Ask how findings are validated. Ask whether the service is mainly automated, mainly manual, or a hybrid, and how that distinction shows up in the report. Ask what a retest looks like. Ask how they handle edge cases such as production testing constraints or LLM application testing, where prompt injection, agent behavior, and model-specific abuse paths require careful handling.

If the answers drift back toward slogans, step back.

## **The report should help three audiences at once**

A pentest report succeeds only when it serves different readers without becoming muddled. Engineers need reproducible detail. Security leaders need risk clarity and prioritization. Auditors and procurement stakeholders need defensible evidence that the work happened and matched the stated scope.

Balancing those audiences is harder than it looks. Too much executive language and the report becomes vague. Too much raw output and it becomes unreadable. The best reports solve this by keeping the writing plain and the technical sections disciplined. They do not hide behind jargon. They explain enough for a non-specialist to follow the significance without diluting the technical truth.

One practical test is whether the report captures attack paths, not just isolated weaknesses. Attack paths matter because real compromise is often cumulative. A low-severity information leak, a token handling mistake, and an over-permissive internal role may combine into material risk. If the report cannot connect those dots where they exist, it is leaving value on the table.

That is especially relevant now that environments span APIs, SaaS integrations, cloud IAM, CI/CD systems, and, increasingly, LLM-backed features. Modern testing often touches more than one layer. Good reporting reflects that complexity without becoming chaotic.

## **The simplest standard: can a competent team act on it?**

Strip away the branding, the portal screenshots, the workflow promises, and the market positioning. What remains is a practical question. Can a competent security and engineering team use this report to understand exposure, fix important weaknesses, verify progress, and explain residual risk to leadership?

If yes, the vendor probably did real work.

If no, the report has failed, regardless of how advanced the offering sounded during the sales process.

That is why unverified claims deserve such skepticism. In the case of the name “TexaTenet,” the only defensible position from the verified context is that the existence of the company or product and the attached offensive

security claims could not be confirmed through reliable sources. Buyers should treat that absence of verification as a signal to demand stronger evidence, not as a reason to lower the bar.

A real pentest report earns trust by being specific, reproducible, and restrained. It tells you what happened, what it means, and what to do next. Anything less is not just poor documentation. It is a security risk in its own right.