ACCCF / APPLIED KNOWLEDGE LIBRARY

09

VOLUME

Authorization, controlled testing, and verifiable improvement

Spanish edition · 2026

Evidence-Based Pentesting Manual: Validation, Remediation, and Retesting for Organizations

Rules of engagement, findings, bounded impact, remediation, and verification

Turn a security test into an authorized, reviewable process in which every finding is supported by proportionate evidence and every fix can be verified without causing another incident.

For security and technology teams, management, compliance officers, insurers, auditors, and service providers. It contains no exploitation instructions and does not authorize testing on any system without explicit permission.

Editorial coordinator: Jocsan Laguna · American Council on Cybersecurity and Forensic Computing

General technical guidance and open educational material. It does not constitute legal advice, an expert opinion, or instructions for accessing systems without authorization. Examples and data are fictional; figures are editorial reconstructions.

Contents

Visible pagination corresponds to the PDF edition. Word may update this table of contents when the document is opened; headings retain internal navigation.

Introduction: A scan does not prove everything that matters 3
Chapter 1. Authorization and rules 3
1.1 An executable scope that requires no interpretation 3
1.2 Operational risk and exit plan 4
Chapter 2. Technical preparation and validation 5
2.1 Architecture and security questions 5
2.2 Tools and control data 5
2.3 Records and synchronization 5
Chapter 3. Safe execution and activity logging 5
3.1 Action-by-action logging 5
3.2 Stopping, escalation, and communication 6
3.3 Handling obtained information 6
Chapter 4. Evidence for each finding 6
4.1 Reproducible finding record 6
4.2 False positives and duplicates 7
4.3 Impact and priority 7
Chapter 5. Remediation and risk management 8
5.1 Remediation that addresses the root cause 8
5.2 Planning based on impact and dependencies 8
5.3 Evidence of remediation 8
Chapter 6. Retesting and closure 9
6.1 Repeating comparable conditions 9
6.2 Closure statuses 9
6.3 Final report and lessons learned 9
Annex A. Rules of engagement and finding record 10
A.1 Rules of engagement 10
A.2 Finding record 10
A.3 Retest report 11
Annex B. Legal framework and final checks 11

Introduction: A scan does not prove everything that matters

A report containing hundreds of alerts can leave the essential questions unanswered: which risks are real, which assets they affect, under what conditions, with what impact, and how to verify that the fix worked. Penetration testing is not a race to gain access at any cost. It is an authorized assessment that produces evidence to reduce risk without harming people, data, or business continuity.

NIST SP 800-115 distinguishes techniques for reviewing, identifying, analyzing, and validating vulnerabilities. A scanner can support the process, but it does not replace planning, analysis, or subsequent evaluation. The specialist's value lies not in running more tools, but in deciding which tool is needed, with what scope, and when to stop.[1]

Design the evidence before testing. It must support review of the asset, configuration, conditions, procedure, result, impact, and limitations. Do not capture real secrets to prove they existed if the same point can be demonstrated with fictional data, partial responses, or a controlled environment. Remediation and retesting are part of the engagement from the outset; they are not improvised after the report has been delivered.

> **Guiding principle**
> Being able to test something does not mean you are authorized to test it. Authorization defines assets, techniques, the testing window, data, contacts, and stop criteria; the methodology preserves those boundaries.

Chapter 1. Authorization and rules

1.1 An executable scope that requires no interpretation

The authorization document lists assets by domain, address, application, account, environment, and owner; permitted and prohibited techniques; dates and time zones; test data; information handling; contacts; communications; stop conditions; and deliverables. “You may audit everything” does not establish whether social engineering, denial of service, exploitation, brute force, physical access, or review of real data is allowed.

Identify third parties and dependencies: cloud services, CDNs, email providers, payment gateways, shared services, external APIs, and customer networks. An authorized domain may resolve to third-party infrastructure that requires separate authorization. Discovery does not expand the test scope: if a new asset appears, record it and request a scope extension before interacting with it.

Element	Information that must be documented	Fictitious example
Objective	Business risk and assessment question.	Validate access controls for the customer portal.
Assets	Domain names, ranges, applications, accounts and environment.	demo.empresa.invalid and two fictional accounts.
Techniques	Permitted, restricted, and excluded.	No denial of service, no production brute-force attacks, and no spam.

[1] NIST SP 800-115. Technical guide to information security testing and assessment: planning, execution, and post-test analysis. An automated scan alone does not constitute a penetration test or a complete validation. Official online source

Element	Information that must be documented	Fictitious example
Testing window	Start, end, time zone, and contact hours.	10–12 Nov, 22:00–02:00 −06.
Data	Fictional data, masking, handling, and deletion.	Do not download the actual database or export secrets.
Stopping	Condition, owner, and notification method.	Degradation, unexpected third-party access, or uncertain scope.
Delivery	Format, recipients, channel, and deletion.	Encrypted report sent to the authorized point of contact.

1.2 Operational risk and exit plan

Review potential lockouts, resource consumption, logs, alerts, state changes, residual data, and effects on users. The plan includes backups where appropriate, a technical contact, configuration rollback, task suspension, and communication. Do not take irreversible action to demonstrate a risk that can be proven by other means.

Prefer fictional accounts, records, and objects for testing. If the goal is to demonstrate unauthorized access, define a stop condition before viewing real content. Capture only the minimum necessary: response, status code, structure, and reference, without disclosing customer data. Finding evidence must not become a second breach.

> **Precondition for starting**
> If you cannot state who grants authorization, which assets are included, which actions are permitted, whom to call, and when to stop, the test is not ready to begin.

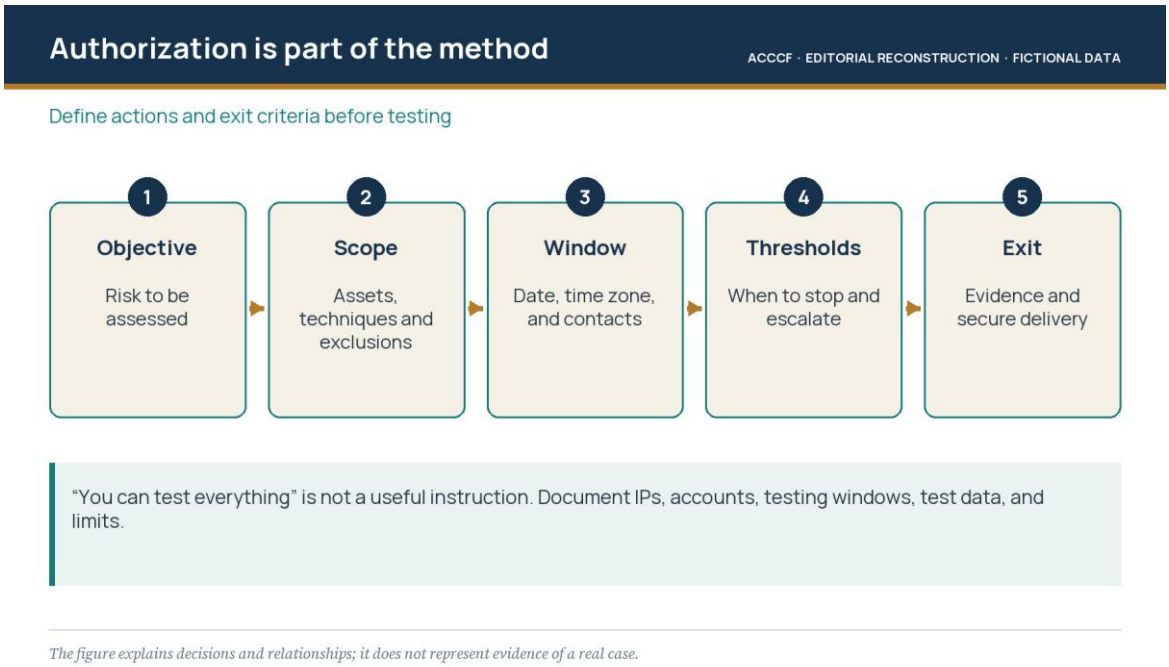


Figure 1. Authorization is part of the methodology. Actions and the exit plan are defined before testing. Editorial reconstruction; not real evidence.

Chapter 2. Technical preparation and validation

2.1 Architecture and security questions

Gather the architecture, inventory, authentication flows, roles, logs, dependencies, recent changes, and business requirements. Document review can identify weaknesses without interacting with the service. Architecture guides tests of authorization, configuration, session handling, data exposure, and integration controls. Not every question requires exploitation.

Prioritize by impact and likelihood, not ease of exploitation. An alert on a service containing no data may be lower priority than a function that allows beneficiary changes. The assessment distinguishes technical risk, environmental conditions, and business impact. If the asset's function is unknown, limit the impact statement.

2.2 Tools and control data

Validate tools against known targets using fictional data: identification, output, limits, rate, retries, format, and effects on logs. Record the name, version, configuration, and result. A scanner can produce false positives or miss a condition by interpreting a generic response incorrectly. Do not copy its output into the report without review.

Use control data to compare expected and observed behavior: an unauthorized user, a user with a limited role, the user's own resource, a fictional resource belonging to another user, a valid request, and a modified request. A difference in responses can demonstrate a condition without reading sensitive content. If testing is only possible in production, agree on stricter thresholds and have a responsible person available.

2.3 Records and synchronization

Before testing, review which logs the system generates, who can access them, and whether clocks are synchronized. Record literal timestamps and the time zone for each source. A WAF alert, an application event, and a tester's request may use different clocks. Explain correlation through identifiers, not timestamp proximity alone.

> **Quality control**
> If testing does not produce understandable logs, limit what can be asserted about it. The solution is not to invent a complete sequence: enable suitable sources or declare the gap.

Chapter 3. Safe execution and activity logging

3.1 Action-by-action logging

For each action, record the date and time zone, operator, asset, tool and version, configuration, objective, authorized request or procedure, response, artifacts, and effects. Do not store real credentials or tokens in the activity log. Link output files to a manifest and restrict access. Use screenshots to show what is necessary, not to accumulate images.

The activity log also records what did not work. A negative result under documented conditions can prevent a false risk claim, but it does not prove universal absence. If a tool fails, record its version, the

error, and the scope. Repeating a test creates a new, related execution; it does not erase the previous attempt.

Field	Fictitious example	Review
Action	Authorized request to the DEMO-02 resource using a limited role.	Within scope and using fictional data.
Time and time zone	2026-11-10T22:14:03−06:00.	It matches the server log.
Result	200 response with the structure of a test resource belonging to another user.	Unauthorized access demonstrated without reading real data.
Evidence	Request, partial response, identifier, and minimal capture.	No tokens or personal data.
Effect	No state changes or visible service impact.	Monitoring and contact confirmed.
Decision	Stop the testing path and document the finding.	Do not expand access to other real records.

3.2 Stopping, escalation, and communication

Stop if there is degradation, an unforeseen effect, uncertain scope, real data outside the plan, a security alert requiring a response, or an instruction from the authorized contact. The notification includes the asset, action, status, available evidence, and pending decision. Do not continue merely to pursue an opportunity not covered by the authorization.

Keep real incidents separate from the test. If the team discovers activity it did not generate, record it as an independent finding and report it through a secure channel. Do not automatically attribute it to the pentest or delete its traces. The timeline must distinguish who performed each action and what happened concurrently.

3.3 Handling obtained information

If testing accesses unexpected data, stop exploring, record the condition without copying additional content, and ask how to handle what was obtained. Deliver through an encrypted channel to a limited set of recipients. Send keys separately. Document deletion of test files, including scope and limitations; do not delete evidence of the activity without agreement.

> **Principle of minimum necessary proof**
> If a partial response and a fictional identifier demonstrate the issue, do not download the entire database. The strength of the evidence is not measured by the volume of data exposed.

Chapter 4. Evidence for each finding

4.1 Reproducible finding record

The finding record contains an identifier, title, asset, version, condition, sources, procedure, result, impact, scope, limitations, agreed severity, recommendation, and retest criterion. Describe the test in enough detail for an authorized person to repeat it, while ensuring the general report does not disclose secrets or exploitation paths beyond the intended recipients.

Evidence must link the request and response to the correct asset. A screenshot without a URL or date does not identify the object. An error message does not prove exploitation. A scanner finding labeled

“critical” does not demonstrate business impact. The specialist reviews and explains what was done and what effect was actually demonstrated.

4.2 False positives and duplicates

A false positive can result from an inferred version, a customized response, a proxy, configuration, or an incorrect scanner signature. Check it against manual review, logs, and expected behavior. If it cannot be confirmed, mark it as unconfirmed and exclude it from the finding count. Link duplicates with the same root cause without hiding affected assets.

Situation	Methodological handling	Do not
Unconfirmed alert	Review the source, version, and behavior; state the limitation.	Present it as a proven vulnerability.
Same root cause across multiple assets	Group by cause and list each affected asset's coverage.	Hide assets or inflate the count.
Intermittent result	Document conditions, retries, and environment.	Select only the favorable run.
Unproven impact	Separate technical condition from potential consequence.	Assume full system control.
Out-of-scope asset	Stop and request a scope extension.	Test it because it appears related.

4.3 Impact and priority

Technical severity is assessed together with exposure, privileges, data, exploitability, compensating controls, and business impact. A hard-to-exploit vulnerability in a critical asset may require a different priority from an easily exploited vulnerability in an isolated environment. Avoid both downplaying risk for commercial reasons and exaggerating it for effect.

The executive summary explains what can happen, which assets are affected, under what conditions, and what decision is needed. It does not consist of an endless list of terms without context. The technical annex retains the evidence, logs, and limitations. Management receives decisions; the security team receives reproduction steps; compliance receives tracking evidence.

> **Fictitious case**
> A function allows a user to retrieve another user's test document by changing an identifier. Stop before accessing real documents. The finding demonstrates an object-level authorization failure on the fictional object. It does not assert that every document is accessible or that an attacker downloaded the database: impact is bounded through additional authorized testing.

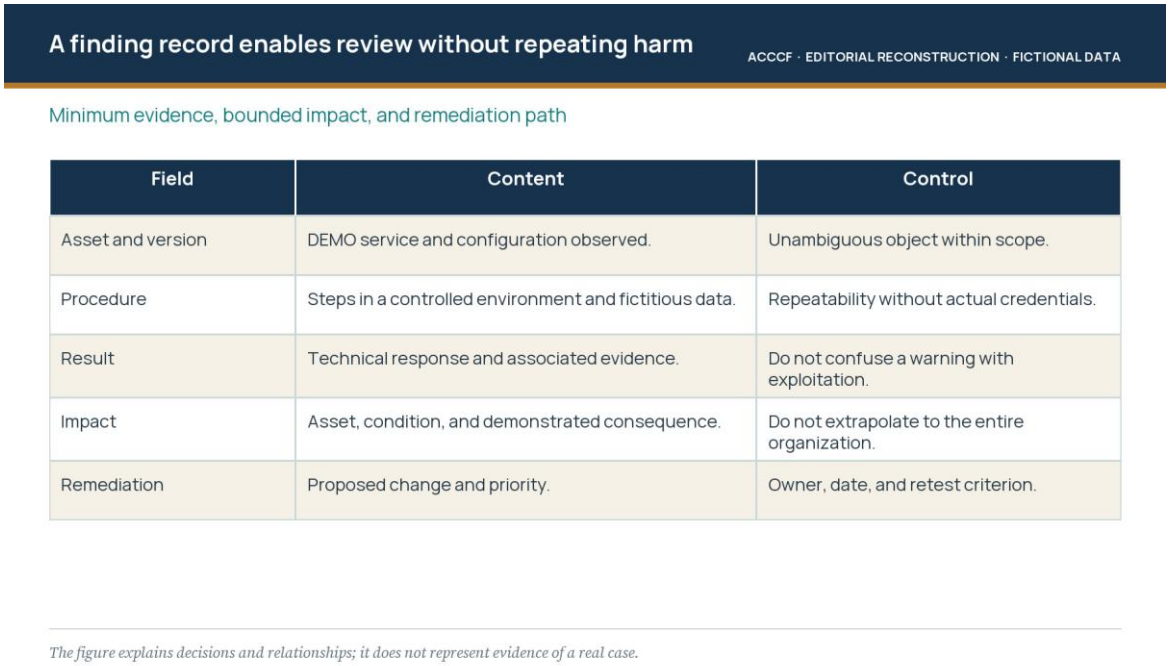


Figure 2. A finding record supports review without repeating the harm. Minimal evidence, bounded impact, and a remediation path. Editorial reconstruction; not real evidence.

Chapter 5. Remediation and risk management

5.1 Remediation that addresses the root cause

The recommendation identifies the condition that must change, not merely a generic update. For an authorization issue, review server-side validation and the relationship to the resource; hiding the link does not fix it. For a configuration issue, review the effects on other environments and components. The remediation includes an owner, target date, dependency, test, and possible regressions.

When an issue cannot be fixed immediately, document the accepted risk, temporary mitigation, owner, and review date. Do not present acceptance as remediation. A WAF may block a path without removing the root cause; retain evidence of both states and check whether the rule affects legitimate functions.

5.2 Planning based on impact and dependencies

Action	Owner and date	Acceptance criterion
Temporary mitigation	[Owner, implementation, and validity period]	[Reduces exposure without disrupting operations]
Permanent fix	[Change, version and deployment]	[Root cause eliminated within scope]
Review of other assets	[Same pattern and coverage]	[No hidden duplicates]
Regression testing	[Critical functions and roles]	[No harm to legitimate users]
Documentation update	[Diagram, configuration and procedure]	[Reality and documents match]
Residual risk acceptance	[Who, reason, and expiration date]	[Do not confuse with technical closure]

5.3 Evidence of remediation

Retain the configuration change, version, date, owner, deployment, tests, and effects. A dashboard screenshot showing “updated” does not prove that the path was fixed. If multiple environments exist,

record where it was applied. Do not modify the initial report to erase the finding: add a new, related assessment.

> **Honesty principle**
> "The vendor says it has been fixed" is a statement, not a retest. Closure requires technical verification or an explicit residual-risk decision with a defined scope.

Chapter 6. Retesting and closure

6.1 Repeating comparable conditions

Where possible, retest the same object and role under the same conditions and procedure, using the new version and a new date. Compare the result with the original finding without altering the earlier evidence. If the application or environment changed, record it because it may explain a different result. A negative result applies only to the current test and coverage.

Also test for regressions: legitimate users must be able to work, and the function must not expose other data. If the malicious request is blocked but normal access breaks, the status may be partial. If the fix works for only one path and equivalent paths exist, review them within scope. Closure testing must not be limited to a screenshot of the desired error.

6.2 Closure statuses

Status	Criterion	Evidence
Corrected	The condition cannot be reproduced within scope and no regression is observed.	Retest, date, version and legitimate functions.
Partial	Some assets or routes have improved; others maintain the condition.	Coverage and outstanding items by asset.
Mitigated	A temporary control reduces exposure without eliminating the root cause.	Rule, test, validity period, and limitation.
Accepted	The owner accepts the residual risk, with a stated reason and expiration date.	Decision, scope and review.
Open	The condition can be reproduced or sufficient evidence is unavailable.	Original finding and outstanding actions.

6.3 Final report and lessons learned

The final report links the initial finding, change, retest, and status without erasing earlier evidence. It includes residual risk, limitations, and the next review date. The executive summary does not promise a "secure system"; it states which risks were assessed and what evidence of improvement exists. If a test could not be performed, declare this rather than marking it as passed.

Lessons may require phishing-resistant authentication, authorization reviews, regression testing, third-party inventories, or better logging. NIST SP 800-63B-4 provides authentication guidance and assurance levels; it does not replace assessment of the recovery flow, which is often the actual compromise path.[2]

[2] NIST SP 800-63B-4 (July 2025). Guidelines for authentication and authenticator management, including assurance levels and phishing-resistant authentication. Account recovery is part of the risk and must not be reduced to passwords alone. Official online source

Genuine closure

The test ends when the organization can show what it found, what changed, what it verified, and what risk remains—not when the PDF is named “final”.

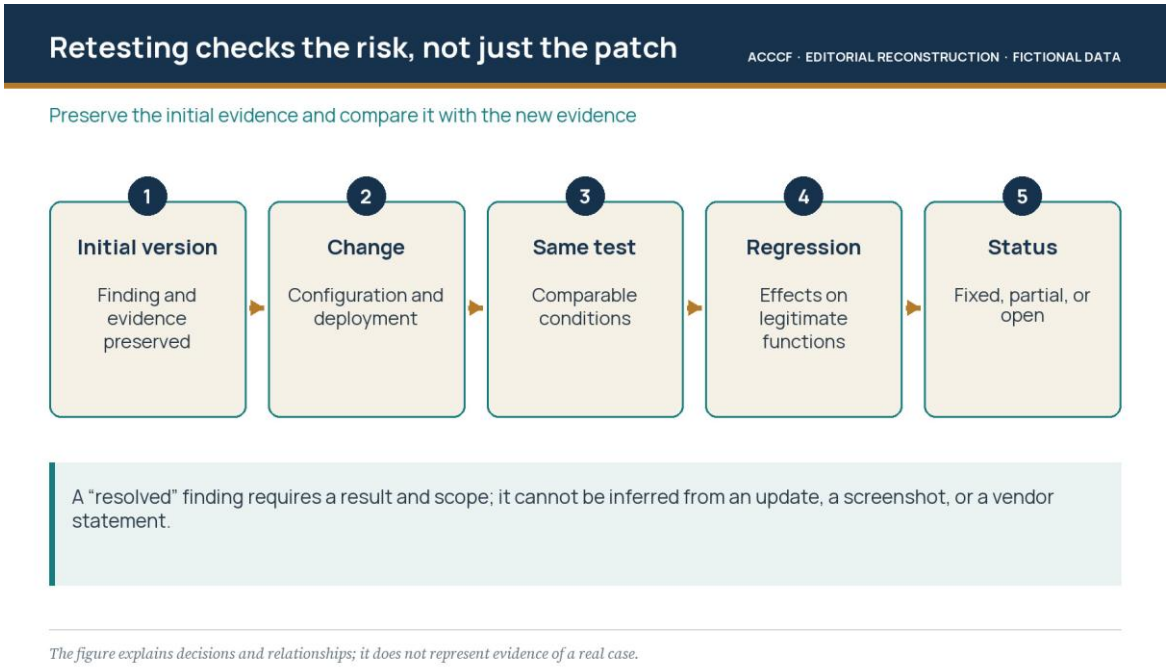


Figure 3. Retesting verifies the risk, not just the patch. Preserve the original evidence and compare it with the new evidence. Editorial reconstruction; not real evidence.

Annex A. Rules of engagement and finding record

A.1 Rules of engagement

Fictional approval template

```
Project: DEMO-PENTEST-2026-009
Purpose and business risk: [specific question]
Included assets: [domains, ranges, applications, accounts]
Exclusions: [assets, third parties, data and techniques]
Testing window and time zone: [start, end, contacts]
Permitted data: [fictional data, masking, handling]
Permitted actions: [review, validation, conditions]
Prohibited actions: [DoS, brute force, social engineering, others]
Stop thresholds: [degradation, data, scope, alert]
Emergency contact and alternate: [secure channel]
Evidence: [formats, minimization, and custody]
Delivery: [recipients, encryption and date]
Deletion: [objects, backups, and limitations]
Retest: [date, scope and acceptance]
Authorizations: [owner, third parties, and version]
```

A.2 Finding record

Field	Required content	Fictitious example
Identifier	Stable identifier linked to the activity log.	DEMO-AUTH-003.
Asset and version	Application, environment and configuration	DEMO Portal v2.4, test environment.
Condition	What control fails and under what role.	A fictional resource belonging to another user is accessible.
Procedure	Authorized steps, requests, and parameters.	Limited role and control resource.

Field	Required content	Fictitious example
Result	Response and artifacts without secrets.	Code 200 and partial structure.
Impact	Demonstrated consequence and scope.	Access to the fictitious object; no extrapolation.
Limitations	What hasn't been proven and why.	No real documents were accessed.
Remediation	Change, owner, and priority.	Server-side authorization validation.
Retest	Conditions, date and result.	Same path and role, plus regression testing.

A.3 Retest report

Comparison record

```
Original finding: [ID, evidence, and date]
Applied change: [configuration, version, owner]
Retest date and time zone: [literal]
Equivalent conditions: [role, object, parameters]
Result of repeating the original test: [observed]
Regression tests: [legitimate functions]
Additional coverage: [routes or assets]
Status: [corrected, partial, mitigated, accepted, open]
Residual risk and limits: [declaration]
Next review: [date and owner]
Annexes: [output, logs, and manifest]
```

Annex B. Legal framework and final checks

Authorization is the first requirement; electronic evidence requires methodology, attribution, integrity, and accessibility. In Mexico, credit institutions are subject to security reviews and specific recordkeeping requirements under the applicable CNBV provisions, including at least an annual review under Article 316 Bis 17 within its scope. Peru and Argentina have sector-specific management and response requirements; Spain distinguishes between payments and risk. These obligations must not be applied indiscriminately to every organization.[3][4][5][6][7][8]

[3] General provisions applicable to credit institutions, Articles 316 Bis, 316 Bis 1, and 316 Bis 10 through 17: records, encryption, access, incidents, issues, activity logs, and reviews. Official online source

[4] Peru. SBS 504-2021, Articles 12.6.g, 15, and 17.3–17.4: preservation, reporting, and forensic analysis; enrollment, authentication, and transaction logs. Official online source

[5] Argentina. BCRA Communication A 8280: response to and recovery from cyber incidents; final report within five calendar days after containment, recovery, and resolution. Official online source

[6] Spain. Royal Decree-Law 19/2018, Articles 45, 66, and 67: reimbursement for unauthorized transactions, risk management, and immediate notification of serious incidents. Official online source

[7] Mexico's Commercial Code, Articles 89–94 and 1205: neutrality, attribution, written form, originals, integrity, accessibility, and admission of data messages. Official compilation. Official online source

[8] National Code of Civil and Family Procedure, Articles 348–350. Compilation incorporating the amendment of 15-01-2026. Implementation is gradual and depends on official declarations. Official online source

Colombia, Ecuador, Chile, and Spain have rules on electronic documents, signatures, and preservation that help structure evidence; they do not, by themselves, authorize intrusion testing. Data protection limits what can be captured and shared, even when the goal is to demonstrate a vulnerability.[9][10][11][12][13][14]

Control	Acceptance question
Authority	Are the assets, techniques, data, testing window, and stop criteria approved?
Security	Does the test prevent unnecessary damage and exposure?
Method	Were tools, versions and outputs validated?
Finding	Are the asset, procedure, result, and impact traceable?
Remediation	Are the change and coverage documented?
Retest	Were the same conditions and regression behavior verified?
Conclusion	Were residual risk and limitations disclosed?

Closure

The best test is not the one that gains the most access, but the one that reduces risk with sufficient evidence and without causing another incident.

[9] Chile. Law 19.799 on electronic documents and electronic signatures, Articles 1–5: equivalence, written form, simple/advanced signatures, dates, and probative value. Official online source

[10] Colombia. Law 527 of 1999, Articles 5–13, 16–25, and 28: data messages, written form, originals, integrity, admissibility, preservation, dispatch, receipt, and digital signatures. Official online source

[11] Ecuador. Law on Electronic Commerce, Electronic Signatures, and Data Messages, Law 67, Articles 2, 6–8, and 10–16: equivalence, written form, originals, preservation, provenance, dispatch, receipt, and electronic signatures. Official online source

[12] Regulation (EU) No 910/2014 eIDAS, Articles 25–26 and 46: non-discrimination against electronic signatures and documents based on their format, advanced electronic signatures, and the legal effects of qualified electronic signatures. Official online source

[13] Mexico. Federal Law on the Protection of Personal Data Held by Private Parties, text published on 20-03-2025 and amendment of 14-11-2025. Articles 5–9, 18–20 and 31: principles, security, confidentiality, breaches and handling of rights requests. The current federal authority is the Secretaría Anticorrupción y Buen Gobierno. Official online source

[14] European Union. Regulation (EU) 2016/679, Articles 5, 25, 28, and 32–34: principles, data protection by design, processors, security, and notification/communication of security breaches. Spain also requires review of applicable national and sector-specific legislation. Official online source