TUTORIAL

# How to Detect Phishing Attacks Using Email Header Analysis

Email headers expose the routing, authentication, and identity signals that help distinguish legitimate messages from phishing. This tutorial shows how to inspect headers, interpret key fields, validate suspicious mail, and operationalize the workflow safely.

**Security - July 21, 2026**

## Why header analysis matters

Phishing often succeeds because the visible sender name and message body look believable. The real evidence is usually in the header trail: the path the message took, which servers handled it, and whether authentication checks such as SPF, DKIM, and DMARC passed. If you can read those signals, you can confirm whether a message was likely sent through an approved mail flow or forged to impersonate a trusted sender.

This tutorial shows how to inspect email headers, identify the fields that matter, and apply a repeatable workflow to decide whether a message is suspicious. By the end, you will be able to extract headers, trace the delivery path, validate sender authenticity, and determine what to verify before using the result in production triage or incident response.

## Prerequisites and stop-here-if checks

## Goal

Make sure the message can be analyzed safely and that the header data is trustworthy enough for a defensible conclusion.

## Action

Collect the original message in a way that preserves full headers and body. Use the message source or raw source, not a forwarded copy pasted into a ticket. If you have access to mail gateway logs, quarantine records, or security tooling, capture the same message ID, timestamp, and recipient details so you can correlate the analysis later.

Check that you have:

- The raw message source, including complete headers
- A safe environment for viewing suspicious content without clicking links or loading remote content
- Basic knowledge of the organization?s legitimate mail providers, gateways, and outbound relays
- Access to DNS or command-line tools for validating SPF, DKIM, and MX records if needed

## Stop-here-if

Stop if you only have a screenshot, a copied body, or a forwarded email without original headers. Those artifacts are not enough to trust the routing path or authentication results. Also stop if the message is tied to an active incident and your environment requires evidence preservation; export the raw source first so later analysis does not alter the evidence chain.

## Expected output

You should have the intact raw message and the minimum context needed to compare the header data with known mail infrastructure.

## Validation

Confirm that the message source contains multiple Received: lines, authentication results, and the sender-related fields. If those are missing, the sample may be incomplete or sanitized.

## Common failure

The most common failure at this stage is inspecting a message that was forwarded or rewrapped by another system. That can replace or hide important header lines and produce false conclusions.

## What to look for in the header set

## Goal

Identify the header fields that provide the strongest evidence of origin, delivery path, and sender identity.

## Action

Focus on the following fields first:

- Received: ? shows each hop the message took through mail servers
- From: ? the display sender seen by the recipient
- Return-Path: ? the envelope sender used for bounce handling
- Reply-To: ? where replies will go if the recipient responds

- Authentication-Results: ? records SPF, DKIM, and DMARC verdicts when a receiving system adds them
- DKIM-Signature: ? indicates a cryptographic signature over selected headers and body
- Message-ID: ? often reveals the sending system?s format and domain patterns
- MIME-Version: and content-type fields ? helpful for spotting malformed or unusual message construction

A legitimate message usually shows a coherent relationship among these values. A phishing message often tries to create visual trust in From: while the routing and authentication fields point somewhere else.

## Expected output

A short list of fields that explain who claimed to send the message, who actually delivered it, and whether the message passed authentication checks.

## Validation

Compare the From: domain to the Return-Path: domain and the domains appearing in Received: lines. They do not need to be identical in every environment, but the pattern should make sense for the organization?s mail flow. If the message claims to come from an internal user but the delivery path begins on an unrelated consumer mail host or foreign infrastructure, treat that as suspicious.

## Common failure

A common mistake is assuming that a matching display name proves legitimacy. Attackers routinely spoof the name while using a different underlying address.

## Step 1: extract the full headers

## Goal

Obtain the full raw header block so you can analyze the message without relying on the mail client?s summary view.

## Action

In most mail clients, open the message source or original message. Save the raw header text and preserve the full content for review. If you need a command-line approach on a local .eml file, start with simple inspection tools:

That prints the header section up to the blank line that separates headers from the body.

If you are analyzing a message exported from a security gateway, keep the original export format intact so you can correlate header lines with any provider-added metadata.

## Expected output

A complete header block, including every Received: line and the authentication fields added by the receiving mail system.

## Validation

Verify that the header block ends at the first blank line. If body content appears mixed into the header output, your extraction method is incomplete.

## Common failure

Mail clients often collapse or hide repeated headers. If you only copy the visible summary pane, you will lose the delivery trail and may miss the strongest evidence.

## Step 2: read the delivery path from bottom to top

### Goal

Reconstruct the message path and identify the first untrusted hop.

### Action

Read the Received: headers from bottom to top. The lowest Received: entry is typically added first, and each later mail server adds another line above it. That means the earliest hop is usually near the bottom of the header block.

Look for:

- The originating host or IP address
- The intermediate relays and whether they belong to known providers
- Timestamp consistency between hops
- Obvious gaps, out-of-order times, or unusual hostnames

If the organization uses How to Implement Zero Trust Network Access in Cyber Security for administrative access paths or remote mail tooling, the same discipline applies here: verify the source, the route, and the policy boundary before trusting the connection.

### Expected output

A rough timeline of how the message moved from sender to recipient, with the earliest external system identified.

## Validation

The chain should be plausible. Large time reversals, missing internal relays, or an initial hop from a host unrelated to your mail providers are strong warning signs. Check whether the IP addresses or hostnames are consistent with your organization?s legitimate mail gateways.

## Common failure

Do not trust a single Received: line in isolation. Attackers can forge header fields above the first receiving system, but they cannot easily fake the complete path once the message has passed through your actual mail infrastructure.

# Step 3: verify sender identity versus envelope identity

## Goal

Determine whether the visible sender matches the underlying mail envelope and the expected domain behavior.

## Action

Compare these values:

- From: for the visible sender identity
- Return-Path: for the bounce address
- Reply-To: for reply handling

A phishing message frequently uses a legitimate-looking From: address while setting Reply-To: to a different domain or abuse account. In some cases, the Return-Path: is a third-party service or a disposable infrastructure domain that has nothing to do with the claimed sender.

Check whether the From: domain is authorized to send mail for the organization or vendor it claims to represent. If the brand is internal, verify that the domain belongs to an approved tenant or mail service and that it matches the expected sender pattern.

## Expected output

A yes/no assessment of whether the visible sender is consistent with the envelope sender and reply destination.

## Validation

If Reply-To: points to a different domain than From:, confirm whether there is a legitimate business reason before trusting the message. Unrelated reply domains are common in phishing and invoice fraud.

## Common failure

A frequent error is assuming the envelope sender is always the same as the display sender. Many legitimate systems use different envelope addresses for bounce management, so you need to compare patterns, not just exact strings.

## Step 4: interpret SPF, DKIM, and DMARC results

## Goal

Use authentication signals to determine whether the message passed domain-level checks that support sender legitimacy.

## Action

Inspect the Authentication-Results: header if present. It usually summarizes the results of SPF, DKIM, and DMARC checks performed by the receiving mail system. Evaluate the results in context:

- SPF checks whether the sending IP was authorized to send mail for the envelope domain
- DKIM checks whether the message was signed and whether the signature remained intact
- DMARC checks whether the authenticated domain aligns with the visible From: domain

A message can still be malicious even if one check passes. For example, a compromised but authorized sender may pass SPF or DKIM while delivering a harmful message. Conversely, forwarded mail or mailing lists can break SPF without indicating phishing.

If you need to validate records manually, use DNS lookups against the domain in question. For example:

That is only a starting point; you still need to interpret the published SPF policy, DKIM selector, and DMARC policy in context.

## Expected output

A clear record of which authentication checks passed, failed, or were absent, plus a judgment on whether the results align with the claimed sender.

## Validation

For a message claiming to be from a protected corporate domain, you should expect Authentication-Results: to show a consistent story. If DMARC fails and the message also contains routing anomalies, the evidence strongly supports phishing or spoofing.

## Common failure

Do not treat SPF or DKIM pass results as proof that the message is safe. Authentication verifies domain control and message integrity, not intent.

# Step 5: spot header inconsistencies that often indicate phishing

## Goal

Identify mismatches that are hard to explain in legitimate mail flows.

## Action

Look for patterns such as:

- A From: domain that does not match the organization being impersonated
- A Reply-To: domain that differs from the claimed sender and routes to an external mailbox
- An unusual Message-ID: domain that does not resemble the sender?s infrastructure
- Missing or malformed DKIM-Signature: when the organization normally signs outbound mail
- Received: lines showing consumer hosting, residential IPs, or unrelated foreign infrastructure
- Authentication results that show spf=fail, dkim=fail, or dmarc=fail

Use caution with isolated anomalies. For example, some legitimate services rewrite headers, add forwarding hops, or omit certain fields. The goal is not to flag every unusual message, but to identify combinations of evidence that do not fit the expected mail path.

## Expected output

A short list of concrete anomalies that justify deeper review or escalation.

## Validation

Ask whether the observed header pattern matches a known, documented sending path. If it does not, and especially if the content pressures the recipient to act quickly, you have a strong phishing indicator.

## Common failure

The most common failure is overfitting on one odd field. A single unusual Message-ID: is weaker evidence than a broken authentication result plus an untrusted first hop plus a suspicious reply domain.

# Step 6: build a repeatable triage workflow

## Goal

Turn header review into a consistent process that analysts and responders can use under time pressure.

## Action

Use a simple decision path:

- Confirm you have the raw message source.
- Check whether the sender and reply domains align with the claimed identity.
- Read the Received: chain from bottom to top.
- Review SPF, DKIM, and DMARC results.
- Compare the header pattern to known good mail flow.
- Escalate if multiple indicators conflict with legitimate delivery.

If your team already has a vulnerability or exposure prioritization workflow, such as How to Detect and Prioritize Critical Vulnerabilities, apply the same principle here: weight the strongest evidence first. A single weak anomaly should not outweigh a broken authentication chain and a suspicious origin.

## Expected output

A documented triage method that produces the same conclusion when different people examine the same message.

## Validation

Run the workflow on a known legitimate internal message and a known phishing sample. The legitimate sample should explain cleanly, and the phishing sample should produce at least one hard mismatch that is visible in the headers.

## Common failure

Teams often jump straight to the body content or link destination. That can be useful, but the header analysis should be the first trust decision because it tells you whether the message is already suspicious before anyone interacts with it.

## Step 7: document findings for incident handling

## Goal

Capture evidence in a form that supports response, blocklist updates, and later investigation.

## Action

Record the following:

- Full raw headers
- Message timestamp and recipient
- Sender, reply, and return-path domains
- Authentication results
- The first untrusted IP or host in the delivery chain
- Any correlation IDs from mail gateways or security tools

Write a concise conclusion that distinguishes evidence from inference. For example: ?The message claims to come from domain A, but the delivery path begins on unrelated infrastructure and DMARC failed; this is consistent with spoofed phishing.? That is more useful than a vague label such as ?suspicious.?

## Expected output

A defensible evidence note that another analyst can review without redoing the entire investigation.

## Validation

Make sure your summary references the exact header fields that support the conclusion. If the conclusion cannot be tied back to the raw source, it is too weak for operational use.

## Common failure

A common failure is documenting only the verdict. Without the supporting header evidence, downstream responders cannot validate the decision or adjust mail controls reliably.

## What a finished, workable state looks like

When this process is working, your team can take a suspicious email, extract the raw headers, trace the delivery path, compare sender and envelope identity, verify SPF/DKIM/DMARC results, and decide whether the message is consistent with legitimate mail flow. The output is not just a yes-or-no answer; it is a repeatable evidence trail that supports triage, user education, and response actions.

In practice, the strongest phishing findings usually come from combinations of evidence: a misleading From: value, a reply address that points elsewhere, an untrusted first hop, and failed authentication. When those signals line up, the message should be treated as malicious until proven otherwise.

The key operational habit is simple: trust the raw headers, validate the path, and document the reasons. That discipline turns email header analysis from an ad hoc review into a reliable detection method.

Use this guidance together with DCSync attack detection and Active Directory permissions audit to connect the workflow with related operational context already available on the site.