**Por qué este análisis de Received: es útil para la adveración**

En adveración de emails, las cabeceras **trace** (especialmente Received:) son la fuente primaria para reconstruir la **ruta efectiva** del mensaje y contrastar su **coherencia temporal y topológica**. El estándar SMTP (RFC 5321) establece que **cada salto** (hop) añade una línea Received, dejando un rastro verificable del tránsito entre MTAs; inconsistencias en orden temporal, continuidad o dominios/IPs pueden apuntar a **manipulación** o errores en la cadena de custodia del mensaje. [ <https://datatracker.ietf.org/doc/html/rfc5321> ]

El módulo construido automatiza tres comprobaciones clave, alineadas con buenas prácticas forenses descritas en la literatura:

1. **Orden cronológico** de hops (fechas tras el último ; normalizadas a UTC),
2. **Continuidad** entre servidores (heurística BY[i+1] == FROM[i], relajada por clúster/organización para infraestructuras modernas como Microsoft 365/Gmail), y
3. **Detección de IPs privadas/loopback** en la cadena. Estas señales están ampliamente aceptadas en trabajos de análisis de cabeceras y guías forenses como indicadores de **spoofing**, rutas reescritas o trayectorias atípicas.

[ <https://commons.erau.edu/jdfsl/vol6/iss2/5/> ]

El uso de IPs **link-local** (fe80::/10) o **loopback** (::1) en determinados hops **no implica fraude por sí mismo**; en entornos Exchange/M365 es habitual observar hops “locales” (MAPI) o direcciones internas dentro del mismo host/cluster. Por eso el módulo **rebaja** la severidad si from==by o si ambos pertenecen a la misma organización/cluster, evitando falsos positivos. (La propia documentación técnica y foros de administración confirman el carácter local de fe80::/10 y la presencia de hops MAPI en O365). [ <https://learn.microsoft.com/en-us/answers/questions/5321462/what-is-the-purpose-of-ipv6-fe80-e> ]

Además, este análisis **complementa** mecanismos de autenticación (SPF, DKIM, DMARC): aunque estos protocolos ayudan a autenticar origen e integridad, **no cubren** todas las formas de reenvío o reescritura; revisar la **coherencia de la ruta** sigue siendo crítico para adveración probatoria y para explicar divergencias cuando la autenticación pasa o falla. NIST y trabajos recientes subrayan el papel **complementario** de SPF/DKIM/DMARC, y la importancia de su **alineamiento** (DMARC) frente a la ruta observable en Received:. [ MobiApps1\_Update, <https://dl.acm.org/doi/fullHtml/10.1145/3471621.3471842> , NIST.TN.1945]

# ¿Qué hace exactamente el código?

* **Parseo robusto de Received:**: normaliza saltos de línea y extrae from, by y la **fecha** (tras el último ;) que convierte a epoch UTC. (RFC 5321 describe estas “trace fields”)
* **Extracción de IPs** con dos capas: primero solo **entre corchetes** (formato canónico); si no hay, **fallback** a IPs en **paréntesis** pegados a from/by (evita confundir horas/IDs con IPs).
* **Clasificación de IPs**: privadas/loopback IPv4 (RFC1918, 127/8, 169.254/16, CGNAT) y privadas IPv6 (::1, fe80::/10, fc00::/7). Se marcan como **INFO** si el hop es local o de la misma organización. (fe80:: es link-local).
* **Continuidad de cadena**: compara BY[i+1] con FROM[i] y, si no coinciden, relaja por **cluster/organización** (p. ej., \*.prod.outlook.com) para infraestructuras distribuidas (O365/Gmail), reduciendo “falsos faltantes”.
* **Salida explicable**: genera un **resumen** (incidencias + detalle por hop) y un **score** simple para priorizar revisión humana; esto se publica en **Analysis Results** junto con los **hashes** de email y adjuntos, facilitando trazabilidad.

# Cómo puntúan los scores

* Partimos de 0 y **sumamos penalizaciones** por hallazgos:  
  **+20** por cada par de hops con **orden temporal inverso**; **+10** por hop **sin fecha**; **+15** por **salto roto** de continuidad (BY[i+1]≠FROM[i] sin pertenecer a mismo cluster/org).
* Para **IPs privadas**: **+15** si parece “expuesta” (no local ni misma org); **+5** si es **local/organizativa** (INFO).
* Rangos orientativos: **0–9** normal; **10–29** revisar (inconsistencias menores o locales); **30–59** sospechoso (rompen continuidad o hay varias señales); **≥60** alta prioridad (orden temporal invertido + otros indicios).
* Interpreta siempre el score junto al **resumen por hop**: una IP fe80:: con from==by es informativa; varios **reverse-order** son mucho más graves.
* El score es **heurístico** y **explicable** (cada punto proviene de una regla concreta), diseñado para **priorizar** y no para decidir en solitario.