

THE AI INSTITUTE / OPERATING INTELLIGENCE

EXECUTIVE CASE NOTE · 10 OCT / 2026



Separate Incident Evidence from Reassurance

After a breach, leaders need a clear account of what is known, what remains uncertain and what people should do next.

Boards, C-suite and senior management

Research period: September 2026 – October 2026

First published: 10 October 2026

Current to 10 October 2026

PUBLICATION RECORD

An Institute thesis and decision framework

An incident update should distinguish confirmed impact, unresolved exposure and the limits of the investigation. Leaders can act before every fact is known without presenting uncertainty as reassurance.



WHAT LEADERS SHOULD TAKE FROM THIS PAPER

- | | |
|----|---|
| 01 | Separate what was accessed from what was copied or changed. |
| 02 | Give every unresolved claim an evidence owner and next review. |
| 03 | Treat AI attribution as a finding to test, not the organising headline. |
| 04 | Keep protective action and notification decisions moving while investigation continues. |

Research method

Original Institute analysis of DIVD’s public case and data-status pages, Zammad’s advisory, a new alert-triage preprint and official regulator guidance. Case facts are attributed to their publishers. The proposed impact register and exercise are Institute recommendations, not an audited account of any organisation’s internal practice.

First published 10 October 2026 · Updated 10 October 2026 · Research period September 2026 – October 2026 · Version 1.0 · Suggested citation: The AI Institute, *Separate Incident Evidence from Reassurance* (2026).

CONTENTS

Inside this paper

01	Confidence is not the same as closure Risk Casebook
02	A public account with unfinished findings What happened
03	An empty result can hide an unfinished search The management condition
04	Connect each claim to a decision The operating practice
05	Share the facts; localise the obligations What travels
06	Rehearse the unfinished update This week’s action
M	Method and limitations
R	References

How to use this brief

Start with the decision and use the framework to challenge the current plan. Figures retain their original populations and definitions. Institute analysis is labelled, and limitations are recorded before the references.

Confidence is not the same as closure

After a breach, executives face two pressures at once: protect the people affected and explain what happened. Neither is served by an early all-clear that the investigation cannot support. The useful management question is not whether the incident has an AI label. It is whether each important statement is grounded in evidence and leads to a proportionate action.

A short impact register can make that discipline practical. For each affected service or information category, record what is confirmed, what remains unresolved, who owns the next check and what people should do meanwhile. Keep it current enough to guide operations and communications from the same account. It should reduce ambiguity, not become another report that nobody uses.

A public account with unfinished findings

On 21 September, an attacker breached the ticketing environment of the Dutch Institute for Vulnerability Disclosure, DIVD. Its case file, updated on 9 October, attributes subsequent activity to AI agents. That is the organisation's assessment of the intrusion, not independent proof of a fully autonomous attack. The case remained open at this review's cutoff. ¹

DIVD's separate data overview distinguishes investigation progress, preliminary findings and final assessment. It reports confirmed loss of email-address data, possible additional exposure and areas still being examined. It also warns that inclusion in its overview does not itself mean a data category was compromised. This is an account from the affected organisation, not a completed independent audit. ²

The supplier's account matters too. Zammad's 5 October advisory says one reported vulnerability is exploitable in older, unsupported versions but not in version 7 and later in practice. It describes the separate privilege-escalation issue as requiring prior server access and says a solution is being worked on. These distinctions do not support a blanket claim that every version has the same remote exposure or that all issues are resolved. ³

An empty result can hide an unfinished search

The condition leaders should address is the gap between a technical observation and the business conclusion drawn from it. Finding no matching records can mean nothing happened. It can also mean the records do not cover the right system, period or activity. A reassuring sentence can erase that distinction before anyone has deliberately decided to accept the remaining uncertainty.

That is the Institute's interpretation, not a finding that DIVD concealed information. Its public separation of investigation stages illustrates a useful alternative. A board should ask what supports a negative finding and what the investigation could not observe. It should not require investigators to prove an impossible universal negative before taking sensible protective action.

New alert-triage research points in the same direction. A preprint testing 1,247 alerts from one multi-stage attack scenario found that unsuccessful evidence searches could contribute to missed attacks. Its proposed review method improved results in that benchmark. This is a controlled study, not a measured failure rate for commercial security teams, and it does not independently explain the DIVD breach.⁴

Connect each claim to a decision

Start with the claims that change someone’s behaviour: whether credentials need replacing, whether a service can reopen, whether customers should take precautions, and whether a notification is required. Attach a source, time window and accountable owner to each. Distinguish information that was reachable from information demonstrably accessed, copied or altered.

Use three plain states: confirmed impact, unresolved exposure, and no signs found within a stated investigation scope. The last is not a guarantee. Record the material gaps beside it, including unavailable supplier records or short retention periods. Ask the incident lead to explain what additional evidence would change the assessment and when the next update is due.

Give communications and customer-support teams the same approved account, with wording suitable for the people affected. Do not let one team promise safety while another is still assessing the same information. Restrict sensitive investigation details to those who need them; transparency does not require publishing exploitable weaknesses or personal information.

If an AI assistant helps draft the update, require a person to check every material statement against the evidence register. Fluent wording can make an unresolved claim sound settled. The accountable incident lead should approve the meaning, not simply the tone.

EXHIBIT / THREE STATES, THREE DIFFERENT MESSAGES

01 Confirmed impact What the evidence establishes and what protection follows. Institute incident practice	02 Unresolved exposure What is unknown, who checks it and what to do meanwhile. Institute incident practice	03 No signs found The scope searched and the limitations that remain. Institute incident practice
--	---	---

Share the facts; localise the obligations

The evidence discipline works across borders, but a single global notification rule does not. The UK Information Commissioner's guidance explicitly accommodates an initial report followed by further information. It also flags that its guidance is under review. Leaders should use current advice for the affected entity and people, rather than copy a deadline from an overseas case.⁵

Singapore's data-protection authority provides a separate framework covering notification criteria, timing and required information. Its guidance should be assessed on its own terms, not treated as equivalent to the UK process. This note does not determine whether a particular incident is legally notifiable.⁶

For North American, European, Asia-Pacific, African, Middle Eastern and Latin American operations, identify the applicable legal and contractual duties with local specialists. Also test practical capacity: can the organisation obtain supplier records, preserve evidence and communicate in the languages people use? This review provides no comparable regional incident-rate data. A resource-constrained team may need external response support; a multinational may need several local notification decisions from one shared factual record.

Rehearse the unfinished update

Run a short exercise using one plausible breach. Give the team incomplete evidence and ask for the first executive update, the protective actions and the next scheduled revision. Include the incident lead, a business owner, privacy or legal support, and customer communications. Look for statements that become more confident as they move between teams.

Success is not a polished declaration that everything is safe. It is a defensible account of what is known, what still needs checking and who will act. That is a more useful basis for trust than certainty delivered before the evidence.

RESEARCH RECORD

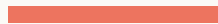
Method and limitations

Method

Original Institute analysis of DIVD's public case and data-status pages, Zammad's advisory, a new alert-triage preprint and official regulator guidance. Case facts are attributed to their publishers. The proposed impact register and exercise are Institute recommendations, not an audited account of any organisation's internal practice.

Limitations

The DIVD investigation remained open. AI attribution is the affected organisation's assessment. Supplier and incident accounts differ in scope and timing; no universal patch or product-safety conclusion is offered. The research is a preprint based on one attack scenario. Regional laws and individual notification duties require current local assessment; this is not legal advice.



Sources and limitations

Source facts link to the original publication. Institute analysis reconciles definitions and turns the material into a management proposition. Frameworks are decision aids, not claims of universal causality. Read every important statistic with its population, date, definition and caveat.

Research cutoff: 10 October 2026. Review cycle: six months, or earlier after a material change in the underlying facts, standards or policy.

REFERENCES

Evidence behind the thesis

01 DIVD: Case DIVD-2026-00014

Incident 21 September 2026; case updated 9 October 2026. Affected organisation's account; investigation open.

02 DIVD: Overview of data investigation

Current overview checked 10 October 2026. Preliminary findings and unresolved categories are not a final forensic assessment.

03 Zammad: Security advisory on the two reported vulnerabilities

5 October 2026. Supplier assessment distinguishes version-dependent remote exposure from local privilege escalation.

04 Tsegai and colleagues: From Investigation Failures to Reliable SOC Agents

New preprint in the 9 October arXiv release. One multi-stage scenario, 1,247 alerts; not a production incident-rate estimate.

05 ICO: UK GDPR data breach reporting

Guidance updated 28 May 2025, checked 10 October 2026; page flags review following the Data (Use and Access) Act. UK-specific.

06 PDPC: Guide on managing and notifying data breaches

Official Singapore guidance; search-index description retrieved, full page browser-challenged. Used only to identify the separate notification framework, not to state deadlines.

THE AI INSTITUTE

Make the decision.
Then measure what changes.

Research · Readiness · Capability

theai.institute