Malice, Mistake or Malfunction?

Detailed research notes | Jane Frankland MBE | 14 November 2026 | Cut-off: 14 November 2026 | v3.1

An independent, AI-assisted investigation of significant, publicly documented digital incidents that began between 1 January and 14 November 2026. This document provides fuller findings, methods, limitations and source notes supporting the accompanying journalistic article. It is not peer-reviewed academic research.

Executive findings

193 incidents were analysed. The reconciled register contained 195 qualifying records, but two precautionary actions were excluded from the cause comparison. Malice was the largest single identified cause, but not a majority. The distinction between data compromise and operational disruption is essential: the former was overwhelmingly malicious, while the latter was frequently unexplained.

Primary cause	Number	Share
Malice	87	45.1%
Mistake	17	8.8%
Malfunction	19	9.8%
External	10	5.2%
Unknown	60	31.1%
Total	193	100%

Different incident types, different patterns

Of 48 data compromises, 44 (91.7%) were malicious, two were mistakes and two unknown. Of 140 operational disruptions, 38 were malicious, 15 mistakes, 19 malfunctions, 10 external and 58 unknown. One further incident qualified as both disruption and compromise, and four through other financial or state-actor criteria. These are mutually exclusive incident classes, so the totals reconcile to 193.

Cause	Operational disruptions (n=140)
Malice	38 (27.1%)
Mistake	15 (10.7%)
Malfunction	19 (13.6%)
External	10 (7.1%)
Unknown	58 (41.4%)

Unknown does not mean non-malicious. If all 58 unknown disruptions were attacks, the malicious share would be 68.6%; if none were, it would be 27.1%. These bounds mean the evidence cannot determine whether attacks or non-malicious causes dominated operational disruption.

Consequences and survivability

For the eight malicious disruptions with sourced durations, median disruption was 252 hours (about 10.5 days), and seven lasted four days or more. Other cause categories had medians between 3.8 and 11.4 hours among cases with durations. However, only eight of 38 malicious disruptions had sourced durations, so this difference cannot be generalised reliably.

Case studies described missing backup paths, shared-provider dependencies and the difficulty of restoring trusted services. Google Cloud London involved a utility failure and failed secondary safeguards; Portugal's national health systems experienced a power-related disruption while a second site was not operational. These cases support qualitative discussion of survivability, not a statistical estimate of how often safeguards fail.

Why the unknown category matters

Among included operational disruptions, 22 of 32 payments and banking cases and 13 of 26 telecom cases lacked a publicly established initiating cause. Regulatory reporting can contain more detailed cause information than public incident announcements, but it concerns different incidents. In particular, the European Supervisory Authorities' 2026 report covers 3,383 major ICT incidents reported by EU financial firms in 2025; the ECB reported that 38% of banks' 2025 major incidents had an IT change as root cause. Neither figure identifies the causes of this investigation's 58 unknown disruptions.

Research design and inclusion

The investigation used 872 structured consequence-first searches (772 sector-region-month cells and 100 supplementary searches) in English and Japanese. Results were grouped into 1,151 organisation clusters, comprising 663 Tier 1 and 488 Tier 2 clusters, all of which underwent AI-assisted verification. Duplicate reports and downstream victims were consolidated to one initiating incident. The incident start date, rather than publication date, determined eligibility.

Fixed inclusion thresholds covered: data concerning 100,000 people (10,000 for sensitive health, biometric or children's data); one-hour interruption affecting 100,000 users or 10 downstream organisations; one-hour interruption of a critical service; any documented safety effect; US$10 million financial loss or a materiality filing; or government-confirmed state-actor activity. A further 337 incidents reported in 2026 had begun before the period and were kept out of the analysis.

Cause definitions and classification

Malice: deliberate initiating act, including cyber intrusion, destructive attack or sabotage.

Mistake: initiating human or organisational action, such as a faulty deployment, configuration change or maintenance step.

Malfunction: unintended technical failure without an identified triggering human change, including latent software, hardware and automation faults.

External: initiating event outside the affected organisation's direct systems, such as an external utility failure.

Unknown: public evidence does not establish an initiating cause; a statement that an event was not an attack does not by itself prove malfunction.

Classification focused on the initiating cause rather than the consequence or a contributing weakness. The register separately records confidence and uncertainty. The two v3.1 late additions—Microsoft 365 / Exchange Online (Unknown) and the FAA nationwide ground stop (Mistake, Low cause confidence)—were based on early accounts and may be revised when official reports appear.

AI assistance and quality controls

AI agents searched, examined sources and proposed incident classifications using a fixed manual. An AI-assisted review process challenged uncertain cases and logged changes. A separate AI model coded a blinded sample of 30 incidents and agreed on 23 (76.7%; Cohen's kappa 0.69). This is AI-to-AI agreement, not independent human validation. No independent human coder validated the full dataset.

Limitations and interpretation

The sample is shaped by public disclosure, media visibility, search languages and significance thresholds. North American and European incidents are more visible. A further 471 records remained provisional and 299 unverified. Recovery duration, safeguards and causes are incompletely documented. The proportions are descriptive of the collected sample, not global prevalence estimates. The large Unknown category can reverse the ranking among operational disruptions.

Source notes and independent comparisons

The principal independent comparators documented in the source verification appendix were: European Supervisory Authorities, report JC 2026 16 (3 June 2026; 2025 ICT incidents); European Central Bank supervisory commentary (24 March 2026; 2025 bank incidents); Ofcom security report (6 March 2026); ENISA telecom security incident reports; Uptime Institute Annual Outage Analysis 2026; IBM Cost of a Data Breach Report 2026; Verizon 2026 Data Breach Investigations Report. These sources have different populations and definitions and were not pooled with the incident register.

Named-incident evidence in the verification appendix includes the Reserve Bank of Australia's incident report; reports on Optus, Bell, Portugal's health systems and Google Cloud London; and preliminary statements concerning the Microsoft and FAA November disruptions. Some named claims rely on secondary reporting. Full incident-level source references are held in the v3.1 research register; this PDF is a readable summary, not a complete source-by-source audit trail.

Versioning and suggested citation

The original v3 dataset was preserved; v3.1 added two late qualifying incidents and documented the changes. Later evidence should be recorded as dated corrections or revisions, not silently inserted into the frozen findings.

Suggested citation: Frankland, J. (2026). *Malice, Mistake or Malfunction? Detailed Research Notes.* Jane Frankland. Dataset v3.1; research cut-off 14 November 2026. Cite the permanent research page on jane-frankland.com once published.