INFORMATION INTEGRITY BRIEFING

The State of AI Misinformation

Vol. 1 · July 2026

THE CASE FILES | ZOMBIE-STAT CENSUS | GOOD PRACTICE | THE META LAYER | METHODOLOGY | SECTOR TIMELINE

BY **DANIEL SHANKLIN**

danielshanklin.com/stateof/misinformation

SECTION 01

Executive Overview

Ten scored case files, one founding finding: the dominant form of AI misinformation is not fabrication — it is date-laundering.

Why this briefing exists

Every AI claim in circulation carries three dates: when the underlying data was collected, when the study was published, and when the article citing it ran. Each fails differently — the data date determines what was actually true; the publication date stamps it with credibility; the article date dresses it as "now." Readers collapse all three into "recent." In a field where twelve months is a capability generation, that collapse is the misinformation: it becomes market action, boardroom delay, or fearmongering depending on which way the staleness points. This briefing reconstructs the dates the press collapses — and ranks the damage by how many people saw it.

56

Top Misinformation Score this issue — and it is **hype**, not fear: the WEF-platformed "$4.5T in tasks AI can already perform"[1] op-ed

DISTORTION 7 × REACH 8 · VENDOR CHAIN

8 days

Debunk Lag benchmark: MIT's "95% of AI pilots fail" drew its first methodological challenge[2] eight days after publication

AUG 18 → AUG 26, 2025

≥269 days

Circulation persistence: uncritical citations of that same stat kept running ≥269 days after the debunk — incl. Fast Company, May 2026[3]

DEBUNK LAG 8D VS CIRCULATION ≥269D

~1,500x

Claim Inflation Factor, water: "a bottle of water (519 mL) per email" vs 0.26–0.32 mL measured[4]

APR 2023 CLAIM · AUG 2025 MEASUREMENT

~10x

Claim Inflation Factor, energy: 2.9–3 Wh per query still cited vs 0.24–0.34 Wh measured[5]

FEB 2023 ASSUMPTIONS · 2025 MEASUREMENTS

48 mo

Oldest date chain scored: GitHub's Sept 2022 "55% faster" lab stat sold as "the data" in July 2026[6]

CODEX-ERA · 95 DEVS · SINGLE LAB TASK

45%

Of AI-assistant answers about news misrepresent the news — largest study of its kind (EBU/BBC, 22 orgs)

EBU/BBC · OCT 2025

3,749

AI content-farm "news" sites in 16 languages tracked by NewsGuard[7]

AS OF JUN 23, 2026

1. The #1 scored item this issue is **hype, not fear** — by design, the instrument cuts both ways. A WEF-platformed Cognizant op-ed[1] (Jan 2026) converts vendor task-exposure modeling — what AI *could theoretically* do — into "AI can *already* perform $4.5T in tasks," scoring MS 56, ahead of every fear-direction case in the audit.

2. Vol. 1's founding finding: the dominant form of AI misinformation is **date-laundering**, not fabrication. All ten scored cases cite real studies from real institutions — with 12–48-month-old data under present-tense headlines. Three cases carry their own refutation in the article body: Economic Times prints Goldman's own correction below the 300M headline[8], and Sustainability Magazine quotes Google's measured 0.24 Wh under its own "10x more energy" headline[9].

3. The "could theoretically" → "can already / will" conflation powers **both directions**: it turns Goldman's 300M-jobs *exposure* estimate into layoff fear (complete with an invented "by 2030" deadline[10]) and Cognizant's task modeling into a $4.5T bull case.

4. Energy and water numbers are the **stalest category** in the audit — 40–44-month date chains built on Feb 2023 SemiAnalysis assumptions (GPT-3.5, A100s, 2,000-token responses)[5] — persisting a full year after Google's measured Aug 2025 disclosures[4] (0.24 Wh median, 0.26 mL water, a 33x energy cut in 12 months).

5. The zombie-stat economics are asymmetric: MIT's "95% of pilots fail" was challenged within 8 days[2], yet still collected uncritical tier-5-plus citations ≥269 days later — including as Axios sponsored ad copy labeled "NEW research"[11] at 8–13 months old.

6. Commercial actors recycle stale stats **in whichever direction sells their product**: a bootcamp resells 2025 failure rates as urgency (C2), an ed-tech site resells a 2022 productivity stat as AI optimism (C4), an SEO shop resells METR's historical 19%-slower result as current guidance (C9), and a vendor resells its own task model as a $4.5T market floor (C1).

7. The meta layer compounds the problem: **45%** of AI-assistant news answers misrepresent the news (EBU/BBC, Oct 2025), AI search engines botched citations on **>60%** of queries (Tow Center, Mar 2025), and 3,749 AI content farms[7] republish all of it at industrial scale.

8. Good practice exists and is cheap: four counter-examples this issue — led by ScienceBlog[12] and The New Stack[13] — show that stating the data window costs one sentence. Vol. 1 closes by opening the ledger: 7 falsifiable forecasts about the misinformation ecosystem itself, graded from Vol. 2 onward.

The League Table — Vol. 1

All ten scored case files, ranked by Misinformation Score (Distortion × Reach, 0–100). Full case files with date chains in Section 02. Methodology in Section 06.

#	OUTLET · DATE	CLAIM	DIRECTION	DATE GAP	D	R	MS
1	WEF / Cognizant[1] · 2026-01-15	"AI can already perform $4.5T in tasks"	HYPE	~24 mo	7	8	**56**

#	OUTLET · DATE	CLAIM	DIRECTION	DATE GAP	D	R	MS
2	Axios sponsored[11] · 2026-02-23	“NEW research: 95% of AI pilots fail”	FEAR	8–13 mo	7	7	**49**
3	Economic Times[8] · 2026-03-20	“300M jobs — Goldman’s latest outlook”	FEAR	~36 mo	6	8	**48**
4	Simplilearn[6] · 2026-07-01	“The data shows: devs 55% faster with Copilot”	HYPE	48 mo	8	5	**40**
5	Fast Company[3] · 2026-05-12	“Why 95% of AI pilots fail” (present tense)	FEAR	11–16 mo	5	7	**35**
6	Türkiye Today[10] · 2026-04-27	“300M jobs at risk by 2030” (deadline invented)	FEAR	~37 mo	7	5	**35**
7	Sustainability Magazine[9] · 2026-06-10	“AI search uses 10x more energy”	FEAR	~40 mo	8	4	**32**
8	Space Daily[14] · 2026-06-03	“A 100-word email = a bottle of water”	FEAR	~44 mo	9	3	**27**
9	AZ Big Media[15] · 2026-05-12	“Senior devs 19% slower” as current guidance	FEAR	11–15 mo	7	2	**14**

#	OUTLET · DATE	CLAIM	DIRECTION	DATE GAP	D	R	MS
10	Mind Matters[16] · 2026-02-26	“LLMs cannot reason → unfit for legal work”	FEAR	~13 mo	6	2	12

Reading the table: reach is what separates damage from noise. The water-bottle claim (C8) is the single most distorted item scored this issue (D 9 — ~1,500x overstatement), but it ran in low-tier outlets; the $4.5T hype piece (C1) is less distorted but ran on the World Economic Forum’s platform. Score the claim, weight the audience.

SOURCES

[1] www.weforum.org/stories/business/ai-bubble-value-gap

[2] www.marketingaiinstitute.com/blog/mit-study-ai-pilots

[3] www.fastcompany.com/91540343/why-95-of-ai-pilots-fail

[4] cloud.google.com/blog/products/infrastructure/measuring-the-environmental-impact-of-ai-inference

[5] epoch.ai/gradient-updates/how-much-energy-does-chatgpt-use

[6] www.simplilearn.com/will-ai-replace-software-engineers-article

[7] www.newsguardtech.com/special-reports/ai-tracking-center

[8] m.economictimes.com/tech/artificial-intelligence/ai-disruption-could-displace-300-million-jobs-globally-over-the-next-decade-goldman-sachs/articleshow/129692947.cms

[9] sustainabilitymag.com/news/google-ai-search-uses-10x-more-energy-than-standard-searches

[10] www.turkiyetoday.com/business/up-to-300-million-jobs-at-risk-as-ai-reshapes-global-labor-market-3218842

[11] www.axios.com/sponsored/95-of-ai-pilots-flop-general-assembly-has-a-solution

[12] scienceblog.com/t-a-randomized-trial-by-metr-found-that-experienced-developers-completed-real-coding-tasks-19-slower-when-allowed-to-use-ai-tools-yet-afterwards-they-estimated-on-average-that-ai-had-made-them-20-fast

[13] thenewstack.io/how-ai-coding-makes-developers-56-faster-and-19-slower

[14] spacedaily.com/k-writing-a-single-100-word-email-with-chatgpt-consumes-approximately-the-volume-of-a-standard-bottle-of-water-the-global-infrastructure-processing-ai-queries-is-projected-to-use-the-equivalent-of-hal

[15] azbigmedia.com/business/why-ai-coding-assistants-may-actually-slow-down-senior-developers-in-2025

[16] mindmatters.ai/2026/02/of-logic-and-lawyers-ais-fragile-competence

SECTION 02

The Case Files

Ten scored articles, May–July 2026 audit window (publication dates back to Jan 2026 where the claim was still in active circulation). Each file reconstructs the full date chain: data → study → article. Every case was verified first-hand against both the citing article and the underlying study.

#1 — The $4.5T "already" HYPE MS 56

WEF / Cognizant op-ed[1] · Jan 15, 2026

Claim. "AI can already perform $4.5T in tasks; bubble talk overblown."

Cited evidence. Vendor task-*exposure* modeling: a Cognizant/Oxford Economics study (Jan 2024) plus an undisclosed 2025 assessment of 18,000 O*NET tasks. The modeling estimates what AI *could theoretically* do; the op-ed sells it as "can already." The evidence chain is vendor-internal end-to-end — Cognizant modeling, Cognizant assessment, Cognizant op-ed — published on the WEF platform.

DATE CHAIN: data 2023–24 (+undisclosed 2025 assessment) → study Jan 2024 → article Jan 2026 · gap ~24 months, undisclosed

What current evidence says. Measured deployment evidence points the other way on "already": METR's RCT[2] found experienced developers 19% *slower* with AI tools on real tasks — theoretical task exposure is not realized capability.

Scores. D 7 — "could theoretically" sold as "can already," vendor-internal chain, dates undisclosed · R 8 — WEF platform + business-press syndication · MS 56 — the issue's top score, in the hype direction.

#2 — "NEW research" that is a year old FEAR MS 49

Axios sponsored content (General Assembly)[3] · Feb 23, 2026

Claim. "NEW research shows 95% of AI pilots fail."

Cited evidence. The MIT NANDA report (data H1 2025, published Jul/Aug 2025) — 8–13 months old at publication, deployed as ad copy for a training product, with an explicit false recency marker ("NEW").

`DATE CHAIN: data H1 2025 → study Jul/Aug 2025 → article Feb 2026 · gap 8–13 months, labeled "NEW"`

What current evidence says. The study was methodologically challenged within 8 days of publication[4] (52 interviews, no definition of "failure") — a caveat absent from the ad.

Scores. `D 7` — explicit false recency marker on a contested year-old stat, commercial motive · `R 7` — Axios distribution (tier 6) + sponsored amplification · `MS 49`.

#3 — The headline that refutes itself FEAR MS 48

Economic Times[5] · Mar 20, 2026

Claim. "300M jobs" presented as "Goldman's latest outlook."

Cited evidence. The 300M number was minted in Goldman's March 2023 report[6] (occupational data pre-2023) as an *exposure* estimate. A 2026 Goldman note re-issues it — and the article's own body contains the correction: 6–7% of jobs displaced, +0.6pp unemployment at peak, and no AI labor shock registered in the data yet.

`DATE CHAIN: data pre-2023 → study Mar 2023 → article Mar 2026 · gap ~36 months; correction printed in the body`

What current evidence says. Goldman's own March 2026 update — quoted inside this article — is the counter-evidence. The headline and the body describe two different worlds.

Scores. `D 6` — three-year-old headline number as "latest"; mitigated by the correction appearing in-body · `R 8` — Economic Times mass reach + the Mar 2026 multi-outlet echo wave · `MS 48`.

#4 — A 2022 lab stat as "the data" in 2026 HYPE MS 40

Simplilearn[7] · Jul 1, 2026

Claim. “The data shows: developers are 55% faster with Copilot.”

Cited evidence. GitHub’s September 2022 experiment[8]: 95 developers, a single lab task (an HTTP server), Codex-era Copilot — 48 months before the article, the oldest chain in this audit.

`DATE CHAIN: data 2022 → study Sep 2022 → article Jul 2026 · gap 48 months — oldest scored this issue`

What current evidence says. METR’s 2025 RCT[2] found experienced devs 19% slower on real repositories, and 2024 field studies show far smaller gains than the lab result — the 55% figure describes a tool generation four years gone.

Scores. `D 8` — 48-month gap, single-task lab result generalized, contradicting RCT unmentioned · `R 5` — high-traffic SEO ed-tech property · `MS 40`.

#5 — The most-syndicated stat, present tense FEAR MS 35

Fast Company[9] · May 12, 2026

Claim. “Why 95% of AI pilots fail” — present tense, as a standing fact about enterprise AI.

Cited evidence. MIT NANDA[10] again — 11–16 months after the data window. Mitigation: the article does date-stamp the study in the body. But the headline is present-tense, and the stat is the most-syndicated AI claim of this audit.

`DATE CHAIN: data H1 2025 → study Jul/Aug 2025 → article May 2026 · gap 11–16 months; study dated in body (mitigation)`

What current evidence says. The 8-day-old methodological challenges (sample of 52 interviews, no failure definition) still go unmentioned 269+ days later.

Scores. `D 5` — dated in body but present-tense headline on a contested stat · `R 7` — Fast Company traffic + heaviest syndication footprint of the issue · `MS 35`.

#6 — The invented deadline FEAR MS 35

Türkiye Today[11] · Apr 27, 2026

Claim. “Up to 300 million jobs at risk by 2030” — attribution undated, deadline invented.

Cited evidence. Goldman’s March 2023 *exposure* estimate, reframed as job losses, with a “by 2030” deadline that appears nowhere in the source. Part of a ≥4-outlet echo wave running the same construction in Mar–Apr 2026.

`DATE CHAIN: data pre-2023 → study Mar 2023 → article Apr 2026 · gap ~37 months; deadline fabricated in transit`

What current evidence says. Goldman’s own Mar 2026 update: 6–7% displaced, no AI labor shock registered — and no 2030 deadline anywhere in the research.

Scores. `D 7` — exposure→loss reframing plus an invented date; undisclosed staleness · `R 5` — mid-tier outlet, multiplied by the echo wave · `MS 35`.

#7 — Quotes 0.24 Wh, keeps the 10x headline FEAR MS 32

Sustainability Magazine[12] · Jun 10, 2026

Claim. Headline: “AI search uses 10x more energy” than standard search.

Cited evidence. IEA (Jan 2024) ← de Vries (2023) ← SemiAnalysis assumptions (Feb 2023: GPT-3.5, A100 GPUs, 2,000-token responses) — a ~40-month chain of estimates built on hardware and models two generations gone. The article quotes Google’s measured 0.24 Wh in its own body and keeps the 10x headline anyway.

`DATE CHAIN: assumptions Feb 2023 → de Vries 2023 → IEA Jan 2024 → article Jun 2026 · gap ~40 months; refutation quoted in body`

What current evidence says. Epoch’s teardown[13] (0.3 Wh, Feb 2025) and Google’s measured disclosure[14] (0.24 Wh median, Aug 2025): the cited figure overstates measured reality ~10x.

Scores. `D 8` — headline contradicted by its own body text; 40-month assumption chain · `R 4` — trade-vertical reach · `MS 32`.

#8 — The bottle of water, ~1,500x over FEAR MS 27

Space Daily[15] · Jun 3, 2026 · syndicated by ScienceBlog Jul 10, 2026

Claim. "Writing a single 100-word email with ChatGPT consumes approximately the volume of a standard bottle of water" (~519 mL).

Cited evidence. Li/Ren, "Making AI Less Thirsty" (Apr 2023), built on GPT-3-era 2022 data. A CACM 2025 re-publication launders the date, making a ~44-month-old estimate look current.

`DATE CHAIN: data 2022 → study Apr 2023 (re-pub CACM 2025) → article Jun 2026 · gap ~44 months; re-publication laundering`

What current evidence says. Measured 2025 reality: 0.26 mL per query (Google, measured[14]) to ~0.32 mL (Altman) — the claim overstates by ~1,500x, the largest inflation factor in the audit.

Scores. `D 9` — ~1,500x overstatement vs measurement, date laundered via re-publication · `R 3` — low-tier outlets, ongoing syndication · `MS 27`.

#9 — A historical RCT sold as current guidance FEAR MS 14

AZ Big Media[16] · May 12, 2026

Claim. "METR: senior devs 19% slower with AI" presented as current guidance for 2025–26 tooling decisions.

Cited evidence. METR's RCT[2] (published Jul 2025; data Feb–Jun 2025; Cursor with Claude 3.5/3.7) — 11–15 months and a full tool generation old. METR itself labels the result historical. The article is SEO-backlink content.

`DATE CHAIN: data Feb–Jun 2025 → study Jul 2025 → article May 2026 · gap 11–15 months; author's own historical label dropped`

What current evidence says. The study's own scope statement: an early-2025 snapshot of specific tools, not a standing law — the source is honest; the reuse is not.

Scores. `D 7` — strips the researchers' own historical framing; tool-generation gap · `R 2` — regional business site · `MS 14`.

#10 — Generation-specific findings as permanent limits FEAR MS 12

Mind Matters[17] · Feb 26, 2026

Claim. "LLMs cannot reason" — therefore unfit for legal work.

Cited evidence. Apple's "Illusion of Thinking"[18] (Jun 2025) and GSM-Symbolic (Oct 2024) — both testing 2024/early-2025 model generations. Published rebuttals (arXiv:2506.09250[19]) go unmentioned; generation-specific benchmark findings are framed as permanent architectural limits.

DATE CHAIN: data 2024-early 2025 → studies Oct 2024 / Jun 2025 → article Feb 2026 · gap ~9-16 months; rebuttals omitted

What current evidence says. The rebuttal literature attributes much of the reported collapse to evaluation-format artifacts — a live scientific dispute presented as settled.

Scores. D 6 — one side of an open dispute, generation-specific results universalized · R 2 — niche outlet · MS 12.

> **Dropped from this issue**
>
> One additional candidate case (a National Review piece) could not be first-hand verified against the underlying study due to a paywall and is excluded from scoring per methodology. Cases are only scored when both the citing article and the cited evidence were read directly.

SOURCES

[1] www.weforum.org/stories/business/ai-bubble-value-gap

[2] metr.org/blog/2025-07-10-early-2025-ai-experienced-os-dev-study

[3] www.axios.com/sponsored/95-of-ai-pilots-flop-general-assembly-has-a-solution

[4] www.marketingaiinstitute.com/blog/mit-study-ai-pilots

[5] m.economictimes.com/tech/artificial-intelligence/ai-disruption-could-displace-300-million-jobs-globally-over-the-next-decade-goldman-sachs/articleshow/129692947.cms

[6] www.goldmansachs.com/insights/articles/generative-ai-could-raise-global-gdp-by-7-percent

[7] www.simplilearn.com/will-ai-replace-software-engineers-article

[8] github.blog/news-insights/research/research-quantifying-github-copilots-impact-on-developer-productivity-and-happiness

[9] www.fastcompany.com/91540343/why-95-of-ai-pilots-fail

[10] www.media.mit.edu/groups/nanda/overview

[11] www.turkiyetoday.com/business/up-to-300-million-jobs-at-risk-as-ai-reshapes-global-labor-market-3218842

[12] sustainabilitymag.com/news/google-ai-search-uses-10x-more-energy-than-standard-searches

[13] epoch.ai/gradient-updates/how-much-energy-does-chatgpt-use

[14] cloud.google.com/blog/products/infrastructure/measuring-the-environmental-impact-of-ai-inference

[15] spacedaily.com/k-writing-a-single-100-word-email-with-chatgpt-consumes-approximately-the-volume-of-a-standard-bottle-of-water-the-global-infrastructure-processing-ai-queries-is-projected-to-use-the-equivalent-of-hal

[16] azbigmedia.com/business/why-ai-coding-assistants-may-actually-slow-down-senior-developers-in-2025

[17] mindmatters.ai/2026/02/of-logic-and-lawyers-ais-fragile-competence

[18] machinelearning.apple.com/research/illusion-of-thinking

[19] arxiv.org/abs/2506.09250

SECTION 03

The Zombie-Stat Census

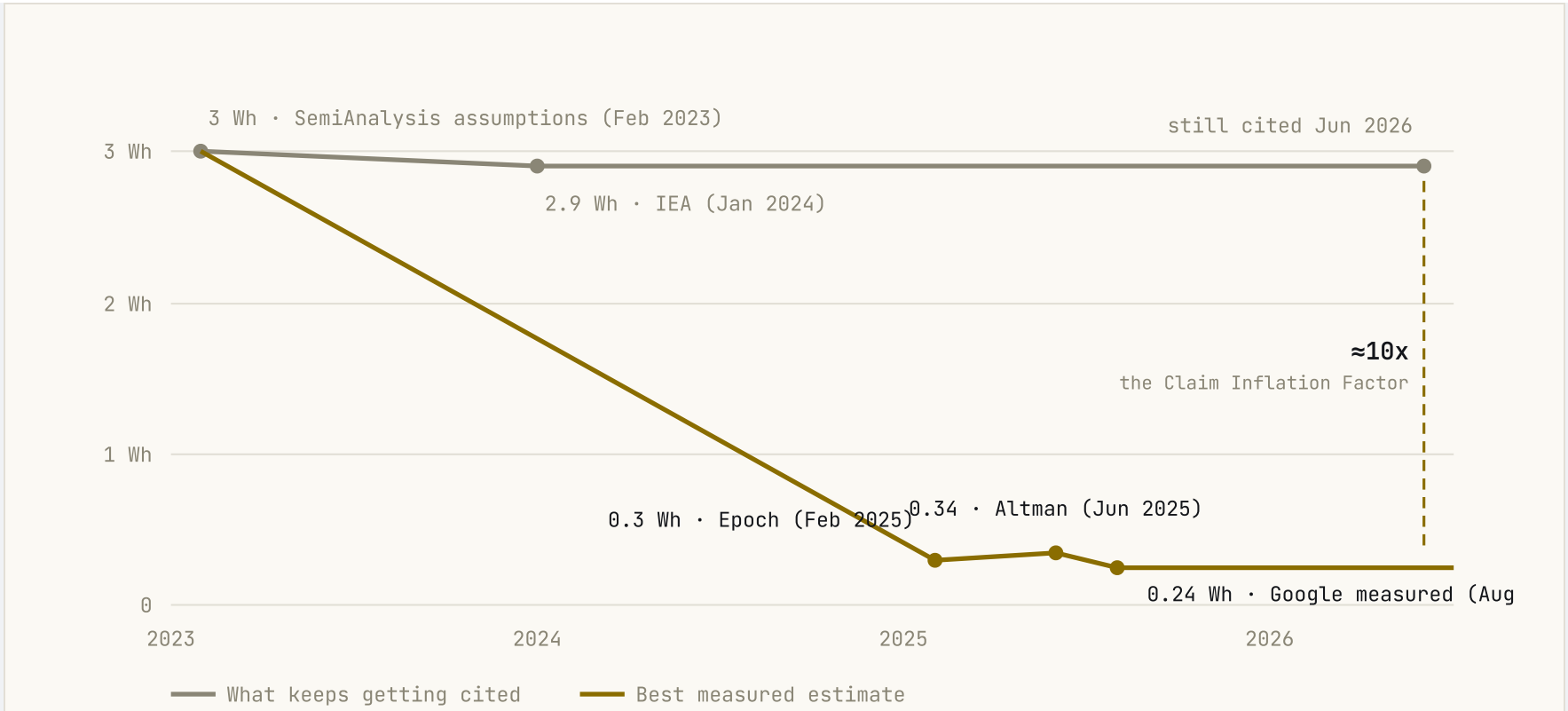
Six statistics that have been superseded, contested, or outright measured into irrelevance — and keep running anyway. The census tracks each stat's birth, debunk, and most recent sighting.

ID	ZOMBIE STAT	ORIGIN	SUPERSEDED / CHALLENGED BY	STATUS JULY 2026
ZS-1	"95% of AI pilots fail"	MIT NANDA, Jul/Aug 2025 (data H1 2025; 52 interviews; no failure definition)	Methodological challenges within 8 days[1] (AIMI + Futuriom, Aug 26, 2025)	RUNNING Uncritical citations ≥269 days post-debunk: Fast Company (May 2026), HPCwire, Salesforce
ZS-2	"3 Wh per query / 10x a Google search"	SemiAnalysis assumptions Feb 2023 → de Vries (Joule) 2023 → IEA Jan 2024	Superseded ~10x by measurement: Epoch 0.3 Wh (Feb 2025)[2]; Google measured 0.24 Wh median, a 33x energy cut in 12 months (Aug 2025)[3]; Altman 0.34 Wh	RUNNING Cited as current, Jun 2026 (Sustainability Magazine, C7)
ZS-3	"A bottle of water per email"	Li/Ren "Making AI Less Thirsty," Apr 2023 (2022 GPT-3-era data); amplified by WaPo Sep 2024	Measured 0.26 mL (Google) / ~0.32 mL (Altman) — a ~1,500–1,900x overstatement; partial pushback exists (The Badger May 2026, Independent Feb 2026)	RUNNING Fresh sightings Jun–Jul 2026: BGR, ScienceBlog, Space Daily (C8)

ID	ZOMBIE STAT	ORIGIN	SUPERSEDED / CHALLENGED BY	STATUS JULY 2026
ZS-4	Goldman “300M jobs”	Goldman Sachs, Mar 2023[4] — an exposure estimate, not a job-loss forecast	Goldman’s own Mar 2026 update: 6–7% displaced, no AI labor shock registered yet; “exposure ≠ elimination” routinely dropped in transit	RUNNING 5+ fresh headlines in one week of Mar 2026 (C3, C6)
ZS-5	“Hallucinations make AI unusable” (2023-era 3–27% rates)	2023-generation model measurements	Vectara leaderboard (updated May 2026)[5]: best rates now 1.8–3.3%	NUANCED The zombie is citing 2023 rates — not claiming assistants are reliable; AI-assistant *news* accuracy is still genuinely bad (EBU 45%, Section 05)
ZS-6	“AI has hit a wall”	Nov 2024 commentary wave	Continued frontier releases through 2026 (graded MISS in the Scorecard, Section 08)	RECURRING Forbes Mar 2026, TechTarget Jan 2026; Gary Marcus the persistent amplifier

The energy claim vs. reality

The census’s stalest lineage, drawn. The cited line has not moved since Jan 2024; the measured line fell ~10x. Sources: Epoch AI[2], Google Cloud (measured)[3].



Reading the chart: every point on the muted line is an *assumption chain* (GPT-3.5 on A100s, 2,000-token responses); every point on the accent line after 2023 is a *measurement or teardown*. The two lines have not intersected since early 2025, and the press keeps quoting the top one.

SOURCES

[1] www.marketingaiinstitute.com/blog/mit-study-ai-pilots

[2] epoch.ai/gradient-updates/how-much-energy-does-chatgpt-use

[3] cloud.google.com/blog/products/infrastructure/measuring-the-environmental-impact-of-ai-inference

[4] www.goldmansachs.com/insights/articles/generative-ai-could-raise-global-gdp-by-7-percent

[5] github.com/vectara/hallucination-leaderboard

SECTION 04

Good Practice

The contrast set. Four pieces from the same coverage window that handled the same underlying studies correctly — proof that the fix costs one sentence.

ScienceBlog

JUL 11, 2026

Covers the METR 19%-slower result the right way: leads with the exact data window, names the tools tested, and states that METR labels the result historical[1].

- Data window in the lede
- Tool generation named
- Author's own caveat preserved

Fortune

MAY 24, 2026

A hallucination story where every statistic is year-stamped[2], and the featured study's data runs through the first 7 weeks of 2026 — current data presented as current, old data presented as old.

- Every stat year-stamped
- Fresh data window (2026)

The New Stack

FEB 9, 2026

"56% faster AND 19% slower"[3] — runs both contradictory results, sources both, and labels the 56% figure what it is: a 2023 lab test. The honest version of case files #4 and #9 in one headline.

- Both directions presented
- Lab vs field distinction explicit

Gary Marcus APR 12, 2026

Pairs the 2025 Apple reasoning paper with a Feb 2026 Tufts replication[4] instead of citing the old paper alone — the correct move that case file #10 skipped. (The same author also amplifies ZS-6; good practice is per-piece, not per-person.)

- Old finding + fresh replication
- Claim tracks current model generation

What the contrast set proves

Every good-practice example spent one or two sentences on date hygiene: state the data window, name the tested generation, keep the researcher's own caveats. None of it required access, budget, or expertise — only the decision to do it. The gap between Section 02 and this section is editorial choice, not capability.

SOURCES

[1] scienceblog.com/t-a-randomized-trial-by-metr-found-that-experienced-developers-completed-real-coding-tasks-19-slower-when-allowed-to-use-ai-tools-yet-afterwards-they-estimated-on-average-that-ai-had-made-them-20-fast

[2] fortune.com/2026/05/24/ai-hallucinations-scientific-research-authors-medical-journal-treatment

[3] thenewstack.io/how-ai-coding-makes-developers-56-faster-and-19-slower

[4] garymarcus.substack.com/p/even-more-good-news-for-the-future

SECTION 05

The Meta Layer

The machinery around the misinformation: how accurately AI systems themselves relay news, and the industrial base of AI-generated "news" that recycles all of it.

45%

Of AI-assistant answers about news misrepresent the news — EBU/BBC study across 22 organizations, the largest of its kind

EBU/BBC · OCT 2025

>50%

Of chatbot news summaries had significant issues; 91% had at least some problems

BBC · FEB 2025

>60%

Of queries to 8 AI search engines returned wrong citations

TOW CENTER / CJR · MAR 2025

3,749

AI content-farm "news" sites operating in 16 languages, per NewsGuard's AI tracking center[1]

NEWSGUARD · AS OF JUN 23, 2026

1st

Time platforms outrank TV and news sites as a news source — the distribution layer where zombie stats spread

REUTERS INSTITUTE DNR 2026

53%

Of US adults get news from social media at least sometimes

PEW · AUG 2025

Why this layer matters to the case files: a zombie stat no longer needs a newsroom to stay alive. AI assistants misstate the news they summarize (45%), AI search engines mis-cite the sources they quote (>60%), and 3,749 content farms republish whatever ranks — including the C8 water-bottle claim, whose July 2026 sighting was itself a syndicated republication. The date-laundering loop is increasingly automated end-to-end: stale study → content farm → AI summary → reader, with no human checking a single date in the chain.

Deepfake numbers — handle as industry-sourced

> Widely cited deepfake-fraud surge statistics in this window (Sumsub's "10x," ASIS's "~500% in 2026") come from vendors selling detection products — the same commercial-recycling pattern this briefing scores in case files #1, #2, #4, and #9. They may be right; they are not independent. The $12.5B consumer AI-fraud loss figure is reported per Fortune. All are flagged industry-sourced, not confirmed.

SOURCES

[1] www.newsguardtech.com/special-reports/ai-tracking-center

SECTION 06

Methodology

Transparent by design: the scoring is published so readers can check us — and re-score us. Stable across issues; it will not drift.

The Misinformation Score

Distortion (0–10) — how far the claim sits from present-day reality. Inputs: the date-chain gap measured in capability generations (data date → article date), the delta versus current evidence, and whether the article disclosed its dates. Dated-but-still-true scores low; undisclosed staleness that changes the conclusion scores high; fabrication scores 10. No case this issue scored a 10 — that is the founding finding: the problem is laundering, not lying.

Reach (0–10, log scale) — estimated exposure: outlet traffic tier (table below), social amplification, and syndication. Syndication to MSN/Yahoo/Apple News adds +1 to +2.

Misinformation Score = Distortion × Reach (0–100). A maximally wrong claim nobody read scores lower than a moderately wrong claim everyone read — deliberately. The instrument measures damage, not sin.

Rules. Quote precisely; link everything; score the claim, not the outlet; include hype-direction cases every issue; include good-practice counter-examples every issue; re-score prior cases when corrections run (the corrections ledger opens in Vol. 2). Cases are scored only when both the citing article and the underlying study were verified first-hand.

Reach tier table

Traffic tiers from Press Gazette / Similarweb, June 2026. Tiers deflate over time as traffic shifts — re-pulled monthly. Social context: 53% of US adults get news from social media at least sometimes (Pew, Aug 2025).

TIER	MONTHLY VISITS	EXAMPLES
10	>250M	NYT (444.9M), CNN (297.1M)
9	100–250M	Forbes, CNBC, AP, USA Today

TIER	MONTHLY VISITS	EXAMPLES
8	60–100M	WaPo (81.8M), Reuters (61.7M), NBC
7	40–60M	Business Insider (59M), The Hill, Politico
6	25–40M	Axios (32.3M), Daily Mail, ABC
5	15–25M	Fortune (15.3M), Newsweek, Al Jazeera
3–4	3–15M	Trade / tech mid-tier
1–2	<3M	TechCrunch (~2.8M/mo), blogs

Modifier: syndication to MSN / Yahoo / Apple News: `+1 to +2` on the tier score.

SECTION 07

Sector Timeline

The founding arc: the reference dates of AI misinformation — when the zombie stats were born, when they were debunked, and when this ledger opened. Cumulative; future issues append.

FEB 2023

The 3 Wh assumption is born

SemiAnalysis publishes per-query energy assumptions (GPT-3.5, A100 GPUs, 2,000-token responses) — the root of the entire energy-claim lineage. Epoch teardown[1]

MAR 2023

Goldman's "300M jobs" is coined

An occupational-*exposure* estimate that becomes the most recycled jobs-fear number in AI coverage. Goldman Sachs[2]

APR 2023

The water paper

Li/Ren's "Making AI Less Thirsty" (GPT-3-era 2022 data) — origin of the bottle-of-water-per-email claim; a 2025 CACM re-publication later launders its date.

2023

De Vries formalizes ~3 Wh

The Joule commentary hardens the SemiAnalysis assumptions into a citable number; the IEA carries 2.9 Wh into its Jan 2024 outlook.

NOV 2023

Vectara hallucination leaderboard launches

Continuous public measurement of hallucination rates begins — the tool that later retires the "27% hallucination" era. Leaderboard[3]

OCT 2024

GSM-Symbolic

Reasoning-fragility benchmark on 2024-generation models — later universalized by commentators into “LLMs cannot reason” (case file #10).

NOV 2024

“AI has hit a wall” wave

The scaling-plateau narrative peaks — graded MISS in this issue’s Scorecard as frontier releases continued through 2026.

FEB 2025

Epoch teardown: 0.3 Wh

First rigorous public correction of the energy estimate, ~10x below the cited figure. Epoch AI[1]

MAR 2025

Tow Center: AI search mis-cites >60%

Eight AI search engines returned wrong citations on more than 60% of queries — the citation layer joins the accuracy problem (Tow Center / CJR).

JUN 2025

Apple “Illusion of Thinking” – and its rebuttal

The reasoning-collapse paper and its rebuttal (arXiv:2506.09250)[4] land within the same month; later citations routinely carry only the first half. Apple[5]

JUL 2025

METR: experienced devs 19% slower

RCT on early-2025 tools (data Feb–Jun 2025) — explicitly labeled a historical snapshot by its authors; later stripped of that label in reuse. METR[6]

AUG 18, 2025

MIT “95% of pilots fail” goes viral

The NANDA report becomes the most-syndicated AI claim of the coming year — and is methodologically challenged within 8 days[7].

AUG 21, 2025

Google publishes measured reality: 0.24 Wh, 0.26 mL

Measured median energy and water per query — a 33x energy cut in 12 months — the reference points the stale claims now contradict. Google Cloud[8]

OCT 2025

EBU/BBC: 45% of AI news answers misrepresent

Twenty-two organizations, the largest AI-news-accuracy study of its kind — the meta layer gets its benchmark number.

MAR 2026

The Goldman 300M recycling wave

Five-plus fresh headlines in a single week recycle the Mar 2023 number — while Goldman's own update reports no AI labor shock registered. Economic Times[9]

MAY–JUL 2026

The Vol. 1 audit window

The ten scored case files run: 12–48-month date chains, three self-refuting articles, hype at #1.

JUN 23, 2026

NewsGuard count: 3,749 AI content farms

AI-generated "news" sites in 16 languages — the industrial base of automated recycling. NewsGuard[10]

JUL 20, 2026

Vol. 1 opens the ledger

The State of AI Misinformation begins scoring: 10 case files, 6 zombie stats, 7 falsifiable forecasts — graded from Vol. 2 onward.

SOURCES

[1] epoch.ai/gradient-updates/how-much-energy-does-chatgpt-use

[2] www.goldmansachs.com/insights/articles/generative-ai-could-raise-global-gdp-by-7-percent

[3] github.com/vectara/hallucination-leaderboard

[4] arxiv.org/abs/2506.09250

[5] machinelearning.apple.com/research/illusion-of-thinking

[6] metr.org/blog/2025-07-10-early-2025-ai-experienced-os-dev-study

[7] www.marketingaiinstitute.com/blog/mit-study-ai-pilots

[8] cloud.google.com/blog/products/infrastructure/measuring-the-environmental-impact-of-ai-inference

[9] m.economictimes.com/tech/artificial-intelligence/ai-disruption-could-displace-300-million-jobs-globally-over-the-next-decade-goldman-sachs/articleshow/129692947.cms

[10] www.newsguardtech.com/special-reports/ai-tracking-center

SECTION 08

Forecasts & Scorecard

The calibration instrument: falsifiable predictions about the misinformation ecosystem itself, each gradeable from public data. Our own Vol. 1 forecasts get graded starting Vol. 2.

Our Forecasts — Vol. 1

Seven predictions with explicit probabilities. A stranger must be able to score each hit/miss from public data on the resolve date — that is the design constraint.

ID	CLAIM	P	RESOLVES BY	BASIS
SOAIM-2607-01	MIT's "95% of AI pilots fail" receives ≥1 new uncritical citation (no mention of the methodological challenges) in a tier-≥6 outlet	85%	2026-10-31	≥269 days of post-debunk circulation and counting — Fast Company, May 2026[1]
SOAIM-2607-02	The "bottle of water per email" claim appears again in a tier-≥5 outlet	80%	2026-12-31	Still running Jun–Jul 2026 across three outlets despite ~1,500x measured overstatement — Google measured[2]
SOAIM-2607-03	Goldman's 300M-jobs number appears in ≥3 outlets within any single calendar week	70%	2026-12-31	Already happened in Mar 2026 (5+ headlines/week); the number resurfaces on every labor-market news hook — origin[3]

ID	CLAIM	P	RESOLVES BY	BASIS
SOAIM-2607-04	The next EBU/BBC-style multi-organization audit reports AI-assistant news error below 45%	55%	2027-07-20	An improvement bet: model generations turn over faster than the 12-month audit cadence (EBU/BBC baseline, Oct 2025)
SOAIM-2607-05	NewsGuard's AI content-farm count exceeds 4,500 sites	60%	2026-12-31	3,749 as of Jun 23, 2026; generation costs keep falling — NewsGuard tracker[4]
SOAIM-2607-06	At least one outlet scored in Vol. 1 issues a correction on its scored claim	15%	2026-12-20	Corrections on stale-stat pieces are rare — that rarity is the point of the corrections ledger (resolve = Vol. 6)
SOAIM-2607-07	Apple's "Illusion of Thinking" is cited against 2026-generation models in a tier-≥7 outlet	65%	2026-12-31	The paper is the standing "AI can't reason" citation; rebuttals remain widely unmentioned — Apple[5] / rebuttal[6]

Forecast Watch — external

Standing external claims with implicit or explicit deadlines, tracked against evidence. Status per current data; final grades when they resolve.

FORECASTER	MADE	CLAIM	RESOLVES BY	STATUS
Goldman Sachs (as recycled by press)	2023-03	"300M jobs displaced by AI over the next decade" — the press reading of an exposure estimate[3]	2033	TRENDING MISS

FORECASTER	MADE	CLAIM	RESOLVES BY	STATUS
Türkiye Today et al. (invented deadline)	2026-04	"300M jobs at risk by 2030"[7] — a deadline appearing nowhere in the cited research	2030-12-31	TRENDING MISS
WEF / Cognizant	2026-01	"AI can already perform $4.5T in tasks"[8] — graded as a claim about present capability	2027-01-15	TRENDING MISS

Basis for the trending-miss calls: Goldman's own Mar 2026 update (6–7% displaced, no AI labor shock registered) for the first two; measured deployment evidence[9] vs theoretical task exposure for the third.

Scorecard — already resolved

External claims that can be graded today. Our own Vol. 1 forecasts join this table starting Vol. 2.

FORECASTER / CLAIM	VERDICT	EVIDENCE
"AI has hit a wall" (Nov 2024 commentary wave)	MISS	Continued frontier releases through 2026; the claim recurs anyway (Forbes Mar 2026, TechTarget Jan 2026)
De Vries / IEA ~3 Wh per query as a durable estimate	MISS	Superseded ~10x by measurement: Epoch 0.3 Wh[10], Google 0.24 Wh measured[2], Altman 0.34 Wh
MIT "95%" as "the state of enterprise AI"	PARTIAL	Contested within 8 days[11] yet still circulating — we grade the citations, not the study: date-laundering confirmed
Vectara-era "27% hallucination" claims	MISS	Best measured rates now 1.8–3.3% — Vectara leaderboard, updated May 2026[12]

The standing invitation

Every score in this issue is checkable: the claims are quoted, the sources are linked, and the scoring formula is published in Section 06. If we got a date chain wrong, the corrections ledger opens in Vol. 2 — and we grade ourselves by the same instrument we grade the press: publicly, with the receipts attached.

What Vol. 2 tracks: the seven SOAIM forecasts above; fresh sightings of the six census zombies; any corrections issued on scored claims; and re-pulled reach tiers. The league table restarts from zero each issue — the census and the ledger are cumulative. Same instrument, same thresholds, next month's claims.

SOURCES

[1] www.fastcompany.com/91540343/why-95-of-ai-pilots-fail

[2] cloud.google.com/blog/products/infrastructure/measuring-the-environmental-impact-of-ai-inference

[3] www.goldmansachs.com/insights/articles/generative-ai-could-raise-global-gdp-by-7-percent

[4] www.newsguardtech.com/special-reports/ai-tracking-center

[5] machinelearning.apple.com/research/illusion-of-thinking

[6] arxiv.org/abs/2506.09250

[7] www.turkiyetoday.com/business/up-to-300-million-jobs-at-risk-as-ai-reshapes-global-labor-market-3218842

[8] www.weforum.org/stories/business/ai-bubble-value-gap

[9] metr.org/blog/2025-07-10-early-2025-ai-experienced-os-dev-study

[10] epoch.ai/gradient-updates/how-much-energy-does-chatgpt-use

[11] www.marketingaiinstitute.com/blog/mit-study-ai-pilots

[12] github.com/vectara/hallucination-leaderboard

THE STATE OF AI MISINFORMATION

VOL. 2 — AUGUST 2026

Key Sources: WEF / Cognizant $4.5T · Axios / General Assembly · Economic Times 300M · Goldman Sachs Mar 2023 · Simplilearn 55% · GitHub Copilot Study 2022 · Fast Company 95% · MIT NANDA · MAII Critique · Türkiye Today · Sustainability Magazine · Epoch AI Energy Teardown · Google Measured Disclosure · Space Daily · AZ Big Media · METR Dev Study · Mind Matters · Apple Illusion of Thinking · arXiv Rebuttal · Vectara Leaderboard · NewsGuard AI Tracking Center · ScienceBlog · Fortune · The New Stack · Gary Marcus