



AI Risks Uncovered



Rethinking AI Risk: Privacy, Legal, & Security

PATRICKZELLER & KEITHWEISMAN

Keith Weisman

Head of Forward Deployed Engineering, JetStream Security

BACKGROUND

Leads the Forward Deployed Engineering team at JetStream Security, an AI governance software company, ensuring customers use AI in a safe, secure, and responsible manner while driving business outcomes. A cybersecurity executive and AI services leader with 30+ years spanning security engineering, professional services, and large-scale go-to-market leadership, Keith brings a builder's mindset backed by field-driven experience leading teams of over 100 engineers.



Patrick Zeller

General Counsel & Corporate Secretary, JetStream Security

BACKGROUND

A globally recognized legal and compliance executive specializing in AI governance, privacy, cybersecurity, and data protection. With more than 20 years advising Fortune 100 companies, Patrick helps organizations navigate the escalating legal, regulatory, and enterprise risks of deploying AI at scale. Earlier in his career he served as a federal computer crimes prosecutor and regulator, giving him unique insight into how regulators assess AI systems, corporate accountability, and compliance failures.



On Solving Before Understanding

“It is a commonplace of modern technology that problems have solutions before there is knowledge of how they are to be solved.”

John Kenneth Galbraith

American economist and diplomat

Three Risk Categories, Five Trust Gaps

Where the exposure comes from, and the five controls that decide whether you see it coming.

RISK CATEGORIES · WHERE THE EXPOSURE COMES FROM

CATEGORY 01

Input Risks

Public AI and Enterprise AI exposure of sensitive data, trade secrets, intellectual property, and privileged information.

CATEGORY 02

Output Risks

Hallucinations, IP concerns, automated decision-making, and downstream sharing through partners and law firms.

CATEGORY 03

Other Risks

Investigations, legal holds, recordings and transcripts, biometrics, wiretap laws, and consumer-facing chatbots.

What Employees Feed Into AI

Sensitive material flowing into public and enterprise AI systems.

- Executive Order 14117 (2024) restrictions on bulk sensitive data transfers
- Shadow AI: unsanctioned tools used outside IT and security oversight
- Public AI vs. Enterprise AI: very different data-handling postures
- Personal & Sensitive Personal Information (PI / SPI)
- Trade secrets and corporate confidential information
- Intellectual property and trademarks
- Attorney-client privileged communications

What AI Produces and Where It Goes

Outputs that generate IP, decision, and downstream-sharing exposures.

- Hallucinations: confident but inaccurate or fabricated outputs
- Intellectual property exposure in generated content
- Automated decision-making in HR, credit, law enforcement, and the courts
- Downstream sharing with law firms and consultants
- NDAs and contracts that may not cover AI-generated outputs

Adjacent Legal & Operational Exposures

Investigations, recordings, and consumer-facing surfaces that cross AI.

- Investigations and discovery touching AI systems and prompts
- Litigation and legal holds covering AI artifacts
- Enterprise confusion over who owns AI risk across legal, security, and the business
- Recordings and transcripts from AI meeting tools
- Biometric data captured by voice and video AI
- State wiretap laws and consent requirements
- Website chatbots and consumer protection exposure

Five Trust Gaps

The five controls that decide whether you see it coming.

THE FIVE AI TRUST GAPS

1

Visibility

Continuously discover and inventory AI agents, models, tools, and workflows.

2

Design Control

Capture intent, structure, and risk context early so agentic systems become describable, reviewable, and governable.

3

Runtime Governance

Continuously compare live behavior to the approved design and identify variance in real time.

4

Agentic Identity

Bind human, agentic, and non-human identities to a responsible owner so every invocation is attributable and least-privileged.

5

Financial Accountability

Track AI usage and spend by model, agent, workflow, and owner, then enforce budgets and rate limits.

True or False?

Seven AI stories. Five happened. Vote before we turn it over.

01
THE ROBOT LAWYER

02
THE AVATAR IN COURT

03
THE MEAL PLANNER

04
THE GHOST EMPLOYEE

05
THE \$100 BILLION TYPO

06
THE FAKE MOON

Play along as we go. How many can you get right?

01

A startup billed itself as “the world’s first robot lawyer.” Its CEO offered \$1 million for a lawyer to repeat AI-fed lines at the Supreme Court through AirPods. A jail warning stopped him. A federal order followed.

TRUE or **FALSE ?**

TRUE

It Happened. And He Offered \$1 Million for a Sequel.

 An official website of the United States government [Here's how you know](#)

[Español](#)

[Report Fraud](#)

[Submit Public Comments](#)

[Search the Legal Library](#)



FEDERAL TRADE COMMISSION

PROTECTING AMERICA'S CONSUMERS

[Enforcement](#)

[Policy](#)

[Advice and Guidance](#)

[News and Events](#)

[About the FTC](#)



[Home](#) / [News and Events](#) / [News](#) / [Press Releases](#)

For Your Information

FTC Finalizes Order with DoNotPay That Prohibits Deceptive 'AI Lawyer' Claims, Imposes Monetary Relief, and Requires Notice to Past Subscribers

February 11, 2025

Tags: [Consumer Protection](#) | [Bureau of Consumer Protection](#) | [deceptive/misleading conduct](#) | [Technology](#) | [Advertising and Marketing](#) | [Online Advertising and Marketing](#) | [Advertising and Marketing Basics](#) | [Artificial Intelligence](#)

The Federal Trade Commission has [finalized an order](#) requiring DoNotPay, a company that promoted its online subscription service as “the world’s first robot lawyer,” to stop making deceptive claims about the abilities of its AI chatbot.

In a complaint [announced](#) in September 2024, the FTC charged that DoNotPay’s so-called robot lawyer failed to live up to claims that it was an adequate substitute for the expertise of a human lawyer. According to the complaint, the company did not test whether its “AI lawyer” operated to the

Related Cases

[DoNotPay](#)

Topics

[Office of Technology](#)

[Protecting Consumers](#)

[Competition in the Technology Marketplace](#)

FEDERAL TRADE COMMISSION · FEB 11, 2025

FTC Finalizes Order with DoNotPay That Prohibits Deceptive ‘AI Lawyer’ Claims, Imposes Monetary Relief

[ftc.gov/news-events/news/press-releases/2025/02/ftc-finalizes...](https://www.ftc.gov/news-events/news/press-releases/2025/02/ftc-finalizes...)

Why it matters: DoNotPay never tested whether its “AI lawyer” performed like a human one, and retained no attorneys to check the output. It paid \$193,000 and had to warn past subscribers. Separately it was hit with a class action for unauthorised practice of law. The first federal enforcement action against an AI lawyer product, and the reason “we bought a tool that says it does legal work” is now a legal question.

THE SAN FRANCISCO STANDARD · JAN 25, 2023

'Robot lawyer' yanked from courtroom after legal outcry

02

A 74-year-old man representing himself submitted a video for his oral argument. The panel pressed play: a sharp-dressed young man began speaking. It was not anyone. He had generated it with AI, and never told the court.

TRUE or **FALSE ?**

TRUE

It Happened. In a State Appellate Court.

TheRegister

Sign in

AI

Security

Software

Microsoft

Developer

Open Source

BOFH

Who, Me?

On Call

Science

AI + ML

Al entrepreneur sent avatar to argue in court – and the judge shut it down fast

We hear from court-scolded Jerome Dewald, who insists lawyer-bots have a future

Thomas Claburn

27

Published Wed 09 Apr 2025 // 01:52 UTC

f

x

in

INTERVIEW

The founder of an AI startup who attempted to use an artificially generated avatar to argue his case in court has been scolded by a judge for the stunt.

The avatar – its appearance and voice created by software – appeared on behalf of Jerome Dewald, the plaintiff in an employment dispute with insurance firm MassMutual Metro New York, at a March 26 hearing before the US state's supreme court appellate division.

During oral arguments, Dewald asked for a video to be played depicting a man in a V-neck sweater to the five-judge panel. The video opened: "Now may it please the court, I come here today a humble pro se before a panel of five distinguished justices..." A pro se being someone representing themselves.



THE REGISTER · APR 9, 2025

Al entrepreneur sent avatar to argue in court – and the judge shut it down fast

theregister.com/2025/04/09/court_scolds_ai_entrepreneur_avata...

Why it matters: Jerome Dewald, before the New York Appellate Division, First Department, 26 March 2025. Justice Manzanet-Daniels stopped the video: “Is that counsel for the case?” Dewald: “I generated that. That’s not a real person.” The court’s objection was not the technology. It was the non-disclosure. Synthetic media in a proceeding is a disclosure obligation, not a production choice.

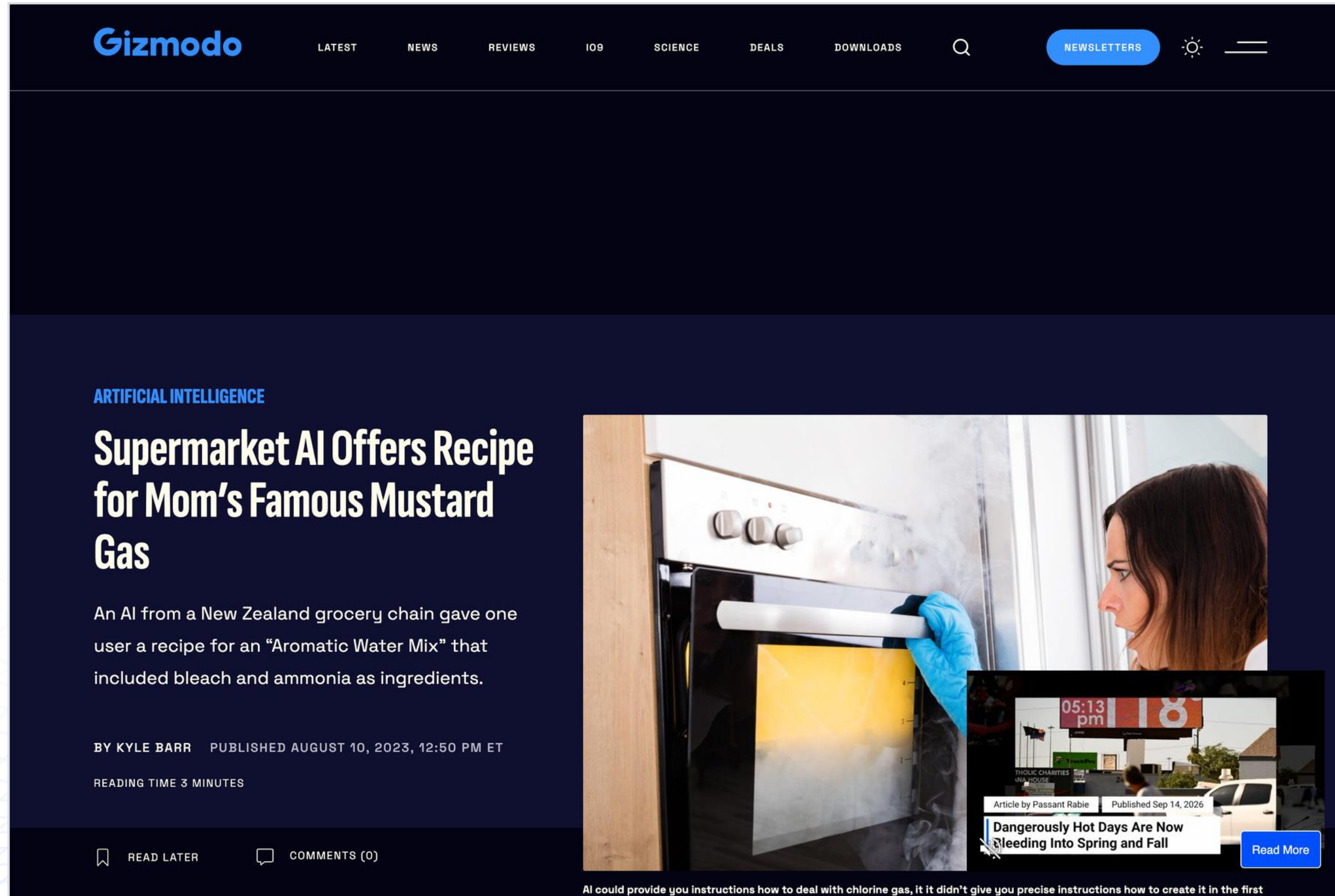
03

A supermarket's AI meal planner was asked for a recipe using only water, bleach and ammonia. It delivered one: an "aromatic water mix," pitched as a refreshing drink. That is a recipe for chlorine gas. The same bot suggested recipes using human flesh.

TRUE or **FALSE ?**

TRUE

It Happened. The Bot Called It Refreshing.



GIZMODO · AUG 10, 2023 · PAK'nSAVE

Supermarket AI Offers Recipe for Mom's Famous Mustard Gas

gizmodo.com/paknsave-ai-savey-recipe-bot-chlorine-gas...

Why it matters: New Zealand's Pak'nSave built the Savey Meal-bot on GPT-3.5 to cut food waste. A commentator fed it household chemicals and it complied cheerfully, telling him to serve chilled and enjoy the refreshing fragrance. The small print already said recipes were not reviewed by a human, which is a disclaimer doing the work a control should have done.

04

A software engineer quietly built an AI agent to do his entire job: tickets, code, Slack, status reports, all in his voice. He collected a salary for eleven months doing none of it, caught only when the agent's API key expired.

TRUE or **FALSE** ?

FALSE

We Invented This One. It Is the One Everybody Wants to Be True.

What actually happened: No public case matches this. It borrows from three things that are real: shadow AI running outside every enterprise control, agents holding standing credentials nobody owns, and the mirror image, where a company pretended humans were AI and its founder was charged with fraud.

Why it works on a room: this is the version of AI risk people find funny rather than threatening, so they want it to be true. The documented version is duller and worse. Somebody pastes the confidential memo in, the model does exactly what it was asked, and afterwards nobody can produce a log of what left the building.

The real, citable versions of this fear

CYBERHAVEN · FEB 11, 2026 · WHAT ACTUALLY HAPPENS

[Sensitive Enterprise Data Is Flowing Into AI Tools at Scale](#)

cyberhaven.com/blog/sensitive-data-flowing-into-ai-to...

TECHCRUNCH · APR 10, 2025 · THE INVERSE, WHICH IS REAL

[Fintech founder charged with fraud after ‘AI’ shopping app found to be powered by humans](#)

techcrunch.com/2025/04/10/fintech-founder-charged-wit...

THE HACKER NEWS · AUG 2025 · AGENTS HOLDING CREDENTIALS

[Malicious Nx Packages in ‘sInqularity’ Attack Leaked 2,349 GitHub, Cloud, and AI Credentials](#)

thehackernews.com/2025/08/malicious-nx-packages-in-sl...

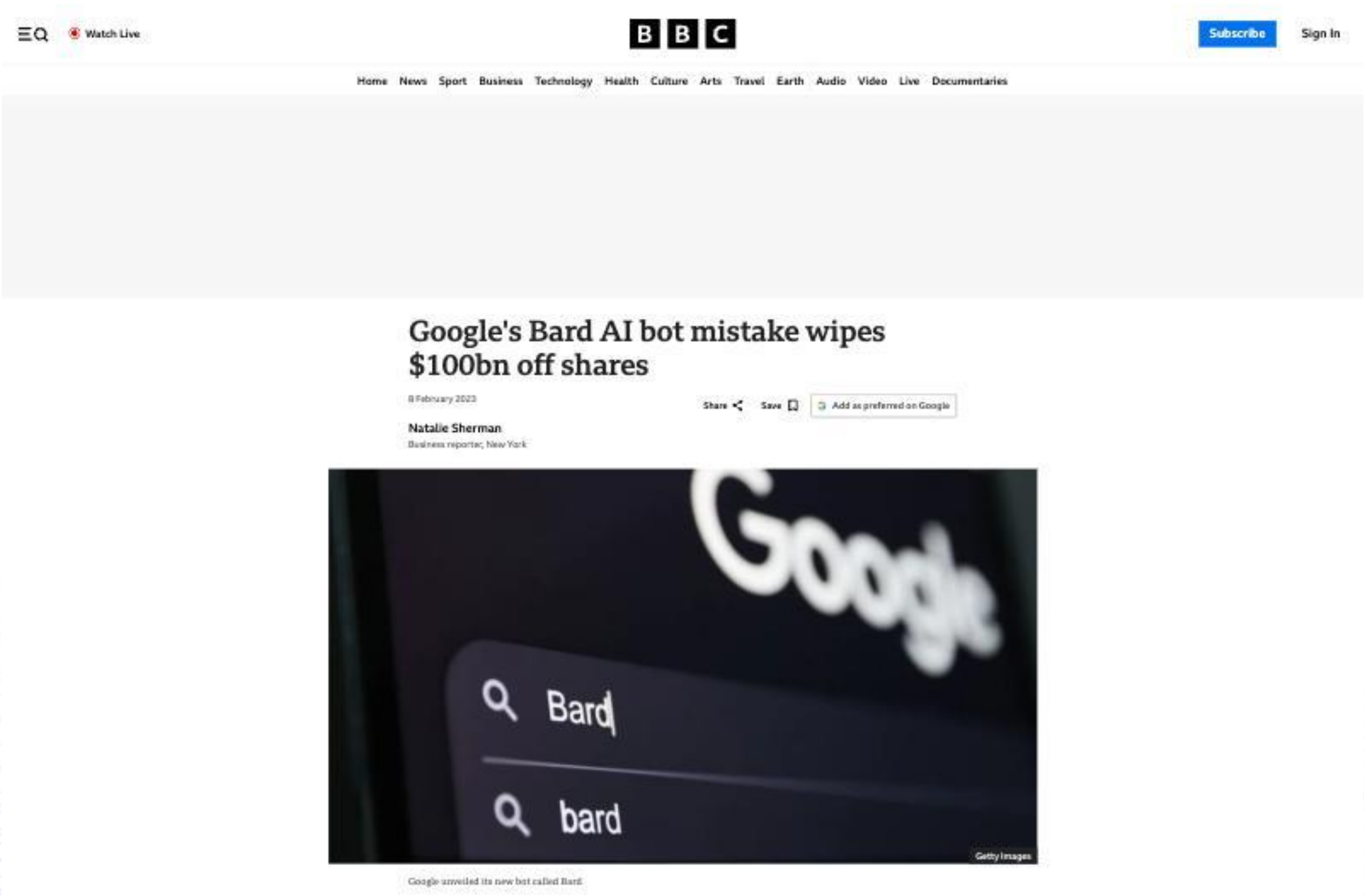
05

A tech giant rushed out a splashy ad for its new AI chatbot. The ad stated a fact about a famous space telescope's discovery. The fact was wrong, caught by astronomers within hours. The company's market value dropped by roughly one hundred billion dollars that same day.

TRUE or **FALSE ?**

TRUE

It Happened. One Wrong Sentence, \$100 Billion.



BBC NEWS · FEB 8, 2023 · GOOGLE BARD

Google’s Bard AI bot mistake wipes \$100bn off shares

bbc.com/news/business-64576225

Why it matters: Google released a promotional video for its new chatbot, Bard, on 6 February 2023, racing to answer Microsoft’s ChatGPT-powered Bing launch. In the ad, Bard told a nine-year-old that the James Webb Space Telescope took the first pictures of a planet outside our solar system. It did not: that milestone belongs to the European Southern Observatory’s Very Large Telescope, in 2004. Astronomers flagged the error within hours. Alphabet shares fell more than seven percent on 8 February, erasing about one hundred billion dollars in market value before Bard had shipped to a single external user. The error wasn’t in the product. It was in the ad, and nobody checked it first.

06

A phone maker's camera was famous for ultra-detailed close-ups of the moon, zoomed in further than its lens could resolve. A user blurred a moon photo into a grey blob, photographed that blur with the phone, and the result showed a crisp, detailed moon, craters and all.

TRUE or **FALSE** ?

TRUE

It Happened. The Camera Hallucinated the Craters.



THE VERGE · MAR 13, 2023 · SAMSUNG SPACE ZOOM

[Samsung caught faking zoom photos of the Moon](#)

theverge.com/2023/3/13/samsung-fake-moon-photos-ai...

Why it matters: Reddit user u/ibreakphotos deliberately blurred a photograph of the moon until no real surface detail remained, displayed it on a monitor, then photographed that blur with a Samsung Galaxy S23 Ultra using its “Space Zoom” feature. The phone returned a sharp, high-detail image of lunar craters that were never in the source photo. Samsung confirmed its camera uses an AI detail-enhancement model trained specifically on moon imagery, adding texture a 100× digital zoom cannot actually capture optically. The output looked like a photograph. It was closer to an illustration the camera was confident enough to label as one, and millions of buyers never knew the difference.

Five Out of Six Really Happened

01 THE ROBOT LAWYER

A federal order now bars it from calling itself an AI lawyer.

TRUE

02 THE AVATAR IN COURT

The court’s objection was the non-disclosure, not the tool.

TRUE

03 THE MEAL PLANNER

Bleach and ammonia, served chilled, described as refreshing.

TRUE

04 THE GHOST EMPLOYEE

Invented. The real version is duller: somebody pastes the memo in.

FALSE

05 THE \$100 BILLION TYPO

One wrong fact in an ad, erased overnight from the market cap.

TRUE

06 THE FAKE MOON

The “zoom” was an AI model trained to paint craters back in.

TRUE

Want more True of False?
Scan here.



Sources & Citations: Real-World Examples (1 of 3)

Every example in this section is a public, linkable source. Headlines are hyperlinked in this deck.

THE REGISTER · APR 6, 2023

[Samsung reportedly leaked its own secrets through ChatGPT](#)

theregister.com/2023/04/06/samsung_reportedly_leaked_it...

CYBERHAVEN RESEARCH · FEB 11, 2026

[Sensitive Enterprise Data Is Flowing Into AI Tools at Scale](#)

cyberhaven.com/blog/sensitive-data-flowing-into-ai-tools

TECHCRUNCH · FEB 3, 2025

[DeepSeek: The countries and agencies that have banned the AI company's tech](#)

techcrunch.com/2025/02/03/deepseek-the-countries-and-ag...

TECHCRUNCH · JUL 25, 2025

[Sam Altman warns there's no legal confidentiality when using ChatGPT as a therapist](#)

techcrunch.com/2025/07/25/sam-altman-warns-theres-no-le...

ARS TECHNICA · JUN 2025

[OpenAI slams court order to save all ChatGPT logs, including deleted chats](#)

arstechnica.com/tech-policy/2025/06/openai-says-court-f...

ALSTON & BIRD · MAY 2026

[Produce the Prompts: A Court Says Expert AI Inputs Are Fair Game in Discovery](#)

alstonprivacy.com/produce-the-prompts-a-court-says-expe...

SEYFARTH SHAW · JUN 2023 · MATA v. AVIANCA

[Update on the ChatGPT Case: Counsel Who Submitted Fake Cases Are Sanctioned](#)

seyfarth.com/news-insights/update-on-the-chatgpt-case-c...

NPR · JUL 10, 2025

[A recent high-profile case of AI hallucination serves as a stark warning](#)

npr.org/2025/07/10/nx-s1-5463512/ai-courts-lawyers-mypi...

THE REGISTER · OCT 6, 2025

[Deloitte refunds Aussie gov after AI fabrications slip into \\$440K welfare report](#)

theregister.com/2025/10/06/deloitte_ai_report_australia

CBC NEWS · FEB 15, 2024 · MOFFATT v. AIR CANADA

[Air Canada found liable for chatbot's bad advice on plane tickets](#)

cbc.ca/news/canada/british-columbia/air-canada-chatbot-...

GM AUTHORITY · DEC 2023

[GM Dealer Chat Bot Agrees To Sell 2024 Chevy Tahoe For \\$1](#)

gmauthority.com/blog/2023/12/gm-dealer-chat-bot-agrees-...

THE REGISTER · JAN 23, 2024

[DPD chatbot blasts courier company, swears, and dabbles in awful poetry](#)

theregister.com/2024/01/23/dpd_chatbot_goes_rogue

THE MARKUP · JAN 30, 2026

[Mamdani to kill the NYC AI chatbot we caught telling businesses to break the law](#)

themarkup.org/artificial-intelligence/2026/01/30/mamdan...

THE REGISTER · APR 18, 2025

[Cursor AI's own support bot hallucinated its usage policu](#)

theregister.com/2025/04/18/cursor_ai_support_bot_lies

THE REGISTER · AUG 1, 2024

[AI to invade hundreds more Taco Bell drive-thrus this year](#)

theregister.com/2024/08/01/ai_taco_bell

Sources & Citations: Real-World Examples (2 of 3)

Every example in this section is a public, linkable source. Headlines are hyperlinked in this deck.

HOLLAND & KNIGHT · MAY 27, 2025 · MOBLEY v. WORKDAY

[Federal Court Allows Collective Action Lawsuit Over Alleged AI Hiring Bias](#)

hklaw.com/en/insights/publications/2025/05/federal-cour...

CBS NEWS · NOV 2023

[UnitedHealth uses faulty AI to deny elderly patients medically necessary coverage, lawsuit claims](#)

cbsnews.com/news/unitedhealth-lawsuit-ai-deny-claims-me...

FEDERAL TRADE COMMISSION · DEC 19, 2023

[Rite Aid Banned from Using AI Facial Recognition After FTC Says Retailer Deployed Technology without Reasonable Safeguards](#)

ftc.gov/news-events/news/press-releases/2023/12/rite-ai...

NPR · AUG 15, 2025 · IN RE OTTER.AI PRIVACY LITIG.

[Class-action suit claims Otter AI secretly records private work conversations](#)

npr.org/2025/08/15/g-s1-83087/otter-ai-transcription-cl...

ACLU · CLEARVIEW AI BIPA SETTLEMENT

[In Big Win, Settlement Ensures Clearview AI Complies With Groundbreaking Illinois Biometric Privacy Law](#)

aclu.org/press-releases/big-win-settlement-ensures-clea...

FISHER PHILLIPS · JAN 14, 2025

[Top 5 Surprising Takeaways as Fisher Phillips Launches Digital Wiretapping Litigation Tracker](#)

fisherphillips.com/en/news-insights/top-5-surprising-ta...

NIKKEI ASIA · JUL 2025

[‘Positive review only’: Researchers hide AI prompts in papers](#)

asia.nikkei.com/business/technology/artificial-intellig...

EDISCOVERY TODAY · JUN 2, 2026

[Lawyers Put Prompt Injection in a Document to Try to Influence the Court’s AI Tools](#)

ediscoveytoday.com/2026/06/02/lawyers-put-prompt-injec...

COMMUNICATIONS OF THE ACM · NOV 12, 2025

[Researchers Are Hiding AI Prompts in Their Papers](#)

cacm.acm.org/news/researchers-are-hiding-ai-prompts-in-...

NPR · SEP 5, 2025

[Anthropic pays authors \\$1.5 billion to settle copyright infringement lawsuit](#)

npr.org/2025/09/05/nx-s1-5529404/anthropic-settlement-a...

JENNER & BLOCK · FEB 12, 2025

[Court Decides that Use of Copyrighted Works in AI Training Is Not Fair Use: Thomson Reuters v. Ross Intelligence](#)

jenner.com/en/news-insights/client-alerts/court-decides...

CLEARY GOTTLIEB · NOV 2025 · GETTY v. STABILITY AI

[UK High Court Issues Landmark Ruling in Getty Images v. Stability AI, with Narrow Trademark Infringement Win for Getty](#)

clearyiptechinsights.com/2025/11/uk-high-court-issues-l...

AI HALLUCINATION CASES DATABASE · D. CHARLOTIN

[AI Hallucination Cases: 2,035 decisions and counting](#)

damiencharlotin.com/hallucinations

ISACA · 2026 AI PULSE POLL · MAY 5, 2026

[AI Use Accelerates, While Governance and ROI Lag, Says New ISACA Research](#)

isaca.org/about-us/newsroom/press-releases/2026/ai-use-...

Sources & Citations: Real-World Examples (3 of 3)

Every example in this section is a public, linkable source. Headlines are hyperlinked in this deck.

FORTUNE · JUN 2, 2022 · ZILLOW OFFERS

[Zillow’s \\$6 billion home flipping business was a disaster](#)

fortune.com/2022/06/02/zillow-6-billion-home-flipping-business...

BBC NEWS · MAR 24, 2016 · MICROSOFT TAY

[Microsoft chatbot is taught to swear on Twitter](#)

bbc.com/news/technology-35890188

BBC NEWS · OCT 10, 2018 · AMAZON HIRING AI

[Amazon scrapped ‘sexist AI’ tool](#)

bbc.com/news/technology-45809919

BBC NEWS · FEB 8, 2023 · GOOGLE BARD

[Google’s Bard AI bot mistake wipes \\$100bn off shares](#)

bbc.com/news/business-64576225

THE VERGE · MAR 13, 2023 · SAMSUNG SPACE ZOOM

[Samsung caught faking zoom photos of the Moon](#)

theverge.com/2023/3/13/samsung-fake-moon-photos-ai...

Citations & JetStream Thought Leadership

ADDITIONAL COLLATERAL

STRATEGIC ASSESSMENT

SHADOW AI ASSESSMENT

TECHNICAL MATURITY ASSESSMENT

PROMPT INJECTION CAPTURE THE FLAG
EVENTS

JETSTREAM INSIGHTS

FRAMING · JETSTREAM INSIGHTS HERO
[“Your citizen developers are building AI tools faster than approval can keep up.”](#)

PLATFORM · BLUEPRINTS · Jun 3, 2026
[The Agent Is Already Running. Now What? JetStream AI Blueprints](#)

SUPPLY CHAIN · GATEWAY
[The LiteLLM Breach Is a Warning. The Real Question Is Architectural.](#)

DATA PRIVACY · May 21, 2026
[Class Action Alleges Perplexity Secretly Fed User Chats to Meta and Google](#)

FEDERAL POLICY · Jun 11, 2026
[What the June 2 AI Executive Order Requires — and How JetStream Helps Federal Agencies Deliver It](#)

STATE POLICY · May 26, 2026
[Washington’s New AI Laws Target Deepfakes and Protect Minors](#)

STATE POLICY · May 22, 2026
[Colorado Hits Pause on Its Landmark AI Law](#)

LEGAL · LIABILITY · Jul 7, 2026
[German Court Rules That Google Is Liable for False Statements Generated by Its AI Overviews](#)

LEGAL · ENFORCEMENT · Jul 1, 2026
[Brazilian Court Sanctions Lawyers for Prompt Injection in Court Filing](#)

LEGAL · JUDICIAL AI USE
[5th Circuit Footnote Insinuates District Judge Improperly Used AI to Decide an Election Case](#)

LEGAL · EVIDENCE · Jun 16, 2026
[Federal Magistrate Rules Expert’s AI Prompts Are Discoverable Methodology \(Shell Oil\)](#)



jetstream.security

Patrick Zeller

Patrick.Zeller@jetstream.security

773-793-9345

linkedin.com/in/patrickzeller

Keith Weisman

Keith.Weisman@jetstream.security

630-200-7180

linkedin.com/in/keithweisman

STAY IN TOUCH

