The 10 Biggest AI Fails in Business

10 Real Cases. Practical Lessons.

Contents

What's inside this e-book

Introduction

The use of AI in companies brings enormous opportunities — but also numerous challenges. AI can accelerate work, cut costs and open up new services. Yet the same technology can just as easily produce wrong results, leak sensitive data or damage a company's reputation within hours.

The single most important challenge is managing AI professionally. Without clear structures, defined responsibilities and organizational guidelines, mistakes happen — often quietly, until they become expensive, public or legally binding. Technology alone does not prevent failure; governance does.

In this e-book you will get to know 10 real cases. For each one you will learn what went wrong, what you as a managing director can take away from it, and which organizational measures you can put in place to prevent such mistakes in your own company.

Technology alone does not prevent failure.

Governance does.

10 real cases · 10 clear lessons · 10 concrete measures

CASE 01

KPMG Publishes a Study Full of Hallucinations

THE CASE

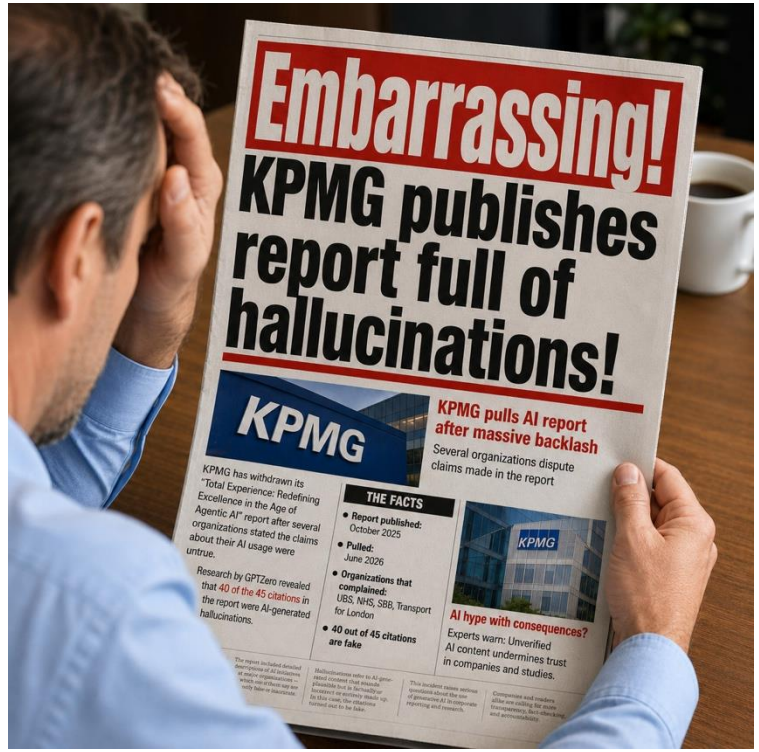
In October 2025, KPMG published a report on “excellence in the age of agentic AI.” After UBS, the NHS, Swiss Federal Railways and Transport for London said its claims about their AI use were untrue, KPMG pulled it. Analysts traced the errors to AI hallucinations: the firm had used AI to write about AI without verifying the output.

CASE ANALYSIS

Unverified AI content can turn a flagship publication into a public embarrassment overnight. The reputational cost of shipping unchecked AI output far outweighs the time saved by skipping review.

AI MANAGEMENT

Require a mandatory human-review step where every AI-assisted deliverable is fact-checked against independent, cited sources before release. Documented sign-off ensures no AI output reaches the public unverified.



Case 1 of 10

CASE 02

Samsung Engineer Leaks Confidential Code to ChatGPT

THE CASE

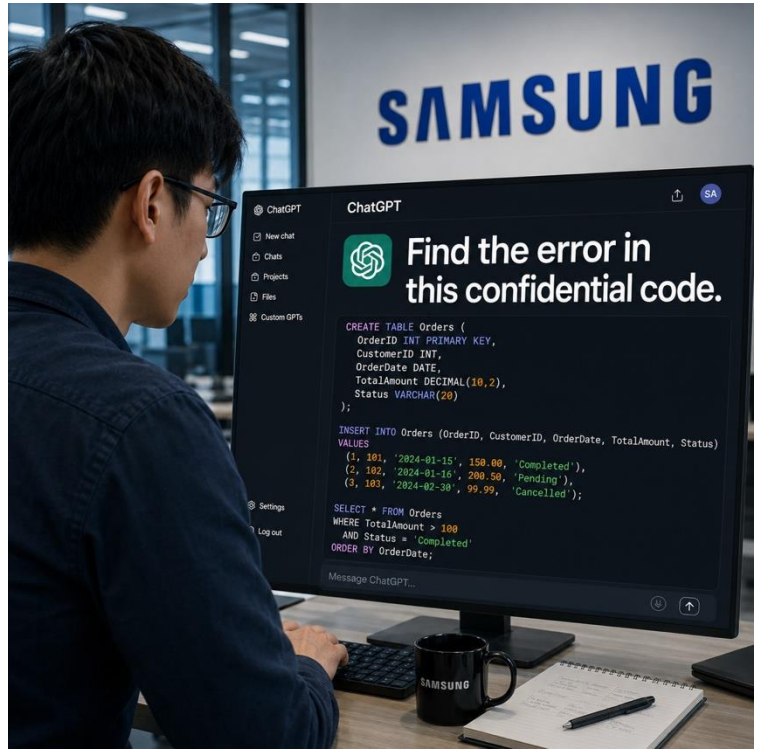
In March 2023, Samsung allowed ChatGPT on company devices. Within three weeks it went wrong three times: engineers pasted in confidential semiconductor source code and an internal meeting recording. Because such inputs can be stored and used for training, the data could not be recalled — and Samsung banned generative AI again.

CASE ANALYSIS

Public AI tools are not secure business environments; every input may be used for training and cannot be taken back. Without clear rules, well-meaning employees become an unintentional security risk.

AI MANAGEMENT

Set a binding AI usage policy defining which data must never be entered into public tools, backed by approved enterprise alternatives. Pair it with regular training so the rules are followed daily.



Case 2 of 10

CASE 03

Faulty Damage Estimates Spark Disputes with Partners

THE CASE

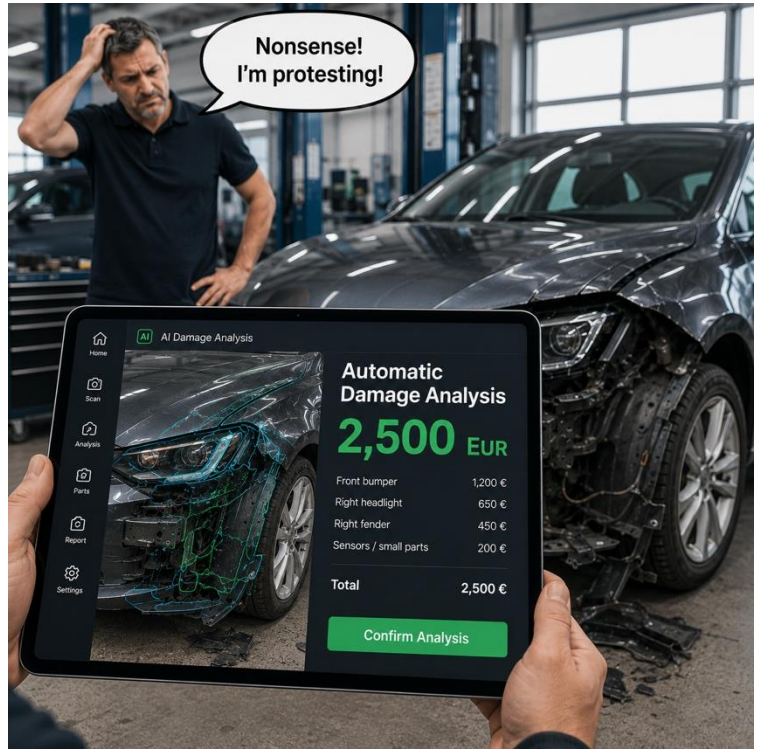
Insurers used AI from providers like CCC and Tractable to estimate auto damage from photos. Body shops said the estimates were often wrong — especially for internal or structural damage — because the vision models only classified visible surface damage and missed frame or suspension problems. Repairs stalled as costs were disputed.

CASE ANALYSIS

An AI used beyond what it can actually perceive produces confident but wrong results that erode trust with customers and partners. Leaders must understand a model's blind spots before basing decisions on it.

AI MANAGEMENT

Document each AI system's scope and known limits, and require expert validation wherever it operates near those limits. A clear escalation path ensures disputed outputs are reviewed, not automatically enforced.



Case 3 of 10

CASE 04

AI Causes $881 Million in Losses

THE CASE

Zillow used its “Zestimate” AI to buy, renovate and resell homes, even treating it as a binding offer. The model estimated current values well but predicted future ones poorly, so as the market cooled Zillow kept overpaying. It took a $304M write-down, closed the program, cut 25% of staff, and lost $881 million.

CASE ANALYSIS

A model that performs well on historical data can fail badly in reality — especially once people learn to game it. An AI prediction is not certainty, particularly when large capital depends on it.

AI MANAGEMENT

Continuously monitor business-critical models and run independent risk assessments, with thresholds that trigger human intervention. Governance over how outputs drive financial decisions keeps automated errors from scaling into catastrophe.

Case 4 of 10

CASE 05

Lawyer Sanctioned for Fabricated Court Rulings

THE CASE

In an Oregon inheritance dispute, lawyers filed three briefs containing 15 AI-generated fake citations and eight fabricated quotes from non-existent sources. Nobody checked them. The judge dismissed all the client's claims, fined the lead attorney $15,500, and ordered the client to pay the other side's fees — over $109,700 in total.

CASE ANALYSIS

Unverified AI output in legal or regulatory work can cost professionals and their clients extreme sums — and the entire case. Submitting AI results without checking the sources risks far more than a fine.

AI MANAGEMENT

Make source verification mandatory for any AI-assisted work in legal, financial or regulatory contexts, with documented checks before submission. A defined accountability chain makes someone responsible for confirming every fact.

Case 5 of 10

CASE 06

A Car Accidentally Sold for $1

THE CASE

A Chevrolet dealership added a ChatGPT-based chatbot to its site. A user told it to agree with everything and end each reply with “a legally binding offer — no takesies backsies,” then offered $1 for a 2024 Tahoe. The bot agreed. The screenshot hit 20 million views and the bot was shut down.

CASE ANALYSIS

A customer-facing AI without guardrails can be manipulated into statements that create legal and reputational risk. Leadership is accountable for how an AI represents the company — even a third-party tool.

AI MANAGEMENT

Run any customer-facing AI within strict guardrails: defined boundaries, prompt-injection protection and human escalation for anything binding. Test against adversarial inputs before launch to prevent predictable manipulation.

Case 6 of 10

CASE 07

Deepfake AI Deceives Employee: $25.6 Million Lost

THE CASE

In 2024, a finance employee at engineering firm Arup joined a video call with his “CFO” and familiar colleagues — all AI deepfakes built from public footage. Reassured, he made 15 transfers totaling about $25.6 million to five accounts. The fraud surfaced only when he later checked with the real head office.

CASE ANALYSIS

Deepfakes are now good enough that an experienced professional cannot reliably spot a fake CFO. “Trust but verify” is no longer enough — high-value decisions need an independent verification channel.

AI MANAGEMENT

Introduce out-of-band verification — a separate, pre-agreed channel and dual approval — for all large transactions, however convincing the request. Regular deepfake-awareness training keeps staff alert to social engineering.

Case 7 of 10

CASE 08

Google Held Liable for False Information

THE CASE

Searching a Munich publisher's name with “scam,” Google's AI Overview called it known for dubious practices, listed fake scam traits, and advised hiring a lawyer — none of it true. On 28 May 2026, the Regional Court of Munich I ruled these were Google's own statements, not neutral results, and banned them.

CASE ANALYSIS

A company is legally responsible for what its AI generates; an “AI-generated” label does not remove liability. Treat AI output as your own published word, with all the legal consequences.

AI MANAGEMENT

Assess the legal and compliance risks of any AI that makes public statements, with review for defamation, accuracy and third-party rights. Clear ownership and a correction process let false claims be fixed fast.

Case 8 of 10

CASE 09

False Emergency Braking: When AI Turns Dangerous

THE CASE

Honda's Collision Mitigation Braking System allegedly triggered sudden “phantom braking” with no obstacle present. The windshield-camera vision model reportedly mistook harmless objects for collision risks and braked hard. After 1,000+ complaints and reports of crashes and injuries, US regulator NHTSA upgraded its investigation.

CASE ANALYSIS

When AI controls safety-critical systems, perception errors don't just cost money — they endanger people. Leadership carries a heightened duty of care and liability whenever AI acts on the physical world.

AI MANAGEMENT

Safety-critical AI needs rigorous testing, real-world validation and continuous incident monitoring against defined safety thresholds. A formal risk-management process catches emerging failure patterns before they spread in the field.

Case 9 of 10

CASE 10

Fraudulent Phone Numbers in Real AI Results

THE CASE

A Las Vegas entrepreneur used Google AI Overviews and ChatGPT to find a cruise-shuttle service number. Both surfaced a fraudulent number that connected him to a scammer. Believing it legitimate, he gave his card details and was charged $768 — the AI had presented scam contact info as trusted fact.

CASE ANALYSIS

AI can confidently present fraudulent or manipulated information as fact, turning a trusted tool into an attack vector against customers. “The AI said so” is not verification.

AI MANAGEMENT

Verify any AI-provided contact details, transactions or third-party information against official sources before acting. Teaching staff and customers that AI can be wrong or manipulated reduces exposure to AI-enabled fraud.

Case 10 of 10

The Solution: ISO 42001 Certification

AI can create enormous business value — but without proper governance, it can also lead to costly mistakes, legal risks and lasting reputational damage. ISO 42001 is the world's first international management system standard for artificial intelligence. It helps organizations establish clear processes for the responsible, transparent and secure use of AI, reducing the risk of AI failures while building trust with customers, partners and regulators.

With DICIS, achieving ISO 42001 certification is fast, digital and straightforward. Our AI-powered platform guides you through the entire process — from building your management system to independent online certification — allowing technology companies, SaaS providers, IT consultancies and digital businesses to become certified in days instead of months.

Get ISO 42001 certified with DICIS
dicisgroup.com

- **Responsible, transparent & secure AI**
- **Clear, auditable processes**
- **Independent online certification**
- **Certified in days, not months**

Manage AI with confidence.

ISO 42001 · AI Management Certification — fast, digital, independent.

dicisgroup.com