



ACF IN ACTION

acfstandard.com

ACF en Action applique l'Agentic Commerce Framework® à des entreprises réelles, à partir de sources publiques. Chaque cas montre qui reste responsable quand l'IA décide.

ACF-CS-028 · Case Study #028 · Édition 2026

Mata v. Avianca (the lawyer and ChatGPT)

Can you present an agent's output without a verifiable trace of its sources?

CARTE D'IDENTITÉ

SECTOR Legal	COUNTRY USA	TECHNOLOGY Research assistant	TYPE Generative AI	AUTONOMY autonomous
CONFIDENCE High				

OUTILS ACF ILLUSTRÉS DANS CE CAS

● ACF-01	● ACF-03	● ACF-06	● ACF-08	● ACF-00	● ACF-02
----------	----------	----------	----------	----------	----------

● mobilisé · ● à prévoir · 6 outils du Toolkit (4 mobilisés, 2 à prévoir, sur 17).

1. Executive summary

In 2023, in a U.S. civil case (Mata v. Avianca), lawyers were preparing a legal brief and delegated case-law research to ChatGPT. The agent fabricated six entirely fake court decisions, with invented case names, citations, and excerpts. The lawyers did not verify them and filed these references with the court. Ordered to produce the texts, they supplied excerpts that were themselves fabricated. On June 22, 2023, the judge sanctioned them with a 5,000-dollar fine.

Read through the lens of ACF, this case is the **"incident"** counterpart to Perplexity (ACF-CS-029): both deal with the same object, the **register and traceability**, by opposite paths. The problem is not merely that the agent hallucinated; hallucination is expected. The governance failure was to present its output **without a register or verification of the sources**. A binding agent assertion has value only if it is traceable back to a real source.

2. Context

Delegating documentary research to a conversational agent is a common and legitimate practice. The difficulty arises when that output is presented to a third party who expects it to be verifiable. A case-law citation only makes sense if it refers to a real, consultable decision. The Mata v. Avianca case shows what happens when this traceability requirement disappears between the agent's output and the filing before the court.

3. Public sources used

- Decision of Judge P. Kevin Castel (U.S. District Court, S.D.N.Y.), June 22, 2023.
- Wikipedia, "Mata v. Avianca, Inc." (sourced facts and decision).
- Seyfarth Shaw, "Update on the ChatGPT Case: Counsel Who Submitted Fake Cases Are Sanctioned" (2023).

4. Reconstructed AI architecture (from public sources)

Case-law research delegated to ChatGPT (Mata v. Avianca case, 2023).

- Nature: general-purpose conversational agent used as a legal research tool.
- Expected function: produce case-law references to support a brief.
- Actual scope observed: the agent produced six **nonexistent** decisions, with **invented** citations and excerpts, presented as real.

The breaking point. The lawyers incorporated these references into the filed brief, without verifying them. When the opposing party pointed out that none of the cited decisions could be found, the court demanded the texts. Instead of acknowledging the error, one lawyer supplied excerpts, also fabricated by the agent. At no point was the chain of assertions grounded in a real, consultable source.

Absent safeguards (what the facts reveal):

- no **verification** of the real existence of the sources before filing;
- no **rule** prohibiting the presentation of an unchecked reference;
- no **register** linking each citation to a real, consultable decision;
- no **human review** required before presenting the output to a third party.

5. ACF mapping

 **Tools used:** ACF-01 (mapping) · ACF-03 (agentic constitution) · ACF-05 (supervision) · ACF-08 (register).

In ACF language, the incident reads as follows:

ACF element	In the Mata v. Avianca case
Named human (accountable)	The lawyer and the firm: it is they who answer for what they present to the court, sources included.
Agent	ChatGPT, used to produce legal references.
Platform	The underlying general-purpose conversational model.
Delegated decision	The case-law research intended for a court filing. Delegating the research does not delegate accountability.
Autonomy level	Autonomous in effect: the output was used without control, as if it had been verified.
Non-delegable zones	Guaranteeing that a source presented to a court exists and is verifiable. This is the boundary that was crossed.
Rule / safeguard	None: neither verification of sources, nor a prohibition on presenting an unchecked reference.
Control third party	After the fact, the court (Judge Castel, S.D.N.Y.).
Traceability	None: no citation referred back to a real decision. There was nothing to trace, therefore nothing to defend.

A schema in ACF graphic language accompanies this case (separate file).

6. Analysis

 **Tools used:** ACF-08 (register) · ACF-03 (verification rule) · ACF-05 (supervision) · ACF-01 (accountability).

What the incident reveals (the governance failure):

- **The real flaw is not the hallucination, it is the absence of a trace.** A general-purpose agent can invent references: that is a known risk. The failure was to present this output to a court **without a register linking each citation to a real source**, and without having checked it before use.
- **Delegating the research does not delegate accountability.** The lawyer answers for what they file, sources included. The agent produces; the human who presents without verifying answers for it. This is the non-delegation of accountability.
- **An assertion without a verifiable source is worthless the moment it binds.** Without a trace leading back to a consultable decision, the output was neither controllable, nor correctable, nor defensible. Later supplying fabricated excerpts only compounded the initial flaw.

What ACF would have made explicit (and enforceable):

- a **register** guaranteeing that each reference is linked to a real, verifiable source, checked before any use;
- a **constitution rule** prohibiting the presentation of an unverified source;
- a **supervision** requiring a human review that verifies the real existence of the sources before filing.

7. ACF toolkit used

This grid indicates the tools that the ACF analysis uses (✓) to read and correct this case, and those to anticipate (→
SOON). It does not mean the company uses ACF: it shows which instruments the framework would bring.

Tool	Status	Role in this case
ACF-08 · Decisions register	✓	Link each reference to a real, verifiable source
ACF-03 · Agentic constitution	✓	Set the rule: never present an unverified source
ACF-05 · Supervision	✓	Require a human review checking the existence of the sources
ACF-01 · Decision map	✓	Locate the accountable party: the lawyer, not the agent
ACF-02 · Criticality matrix	→ SOON	Rank by stake what requires verification before use
ACF-06 · Kill switch	→ SOON	Block a filing until the sources are verified

8. Governance questions (what an AI committee should ask itself)

1. Does an agent assertion without a verifiable source have any value the moment it binds before a third party?
2. Who bears the accountability: the agent that hallucinates, or the one who presents without verifying?
3. Do we have a register linking each agent output to a real, consultable source?
4. Does a human review verify the real existence of the sources before any binding use?
5. Do our rules explicitly prohibit presenting an unchecked source?

9. General lessons

An agent did not fail because it hallucinated: it hallucinates by nature. The governance failure was to use its output without a trace or verification, then to present it to a court. Delegating the research relieves no one of the accountability for what is filed. What was missing was not a better prompt, but a register and a systematic control of the sources.

The transferable lesson: a binding agent output is worth something only if it traces back to a verifiable source. The register is cheaper than the sanction.

10. What ACF would do (reference design)

Prescriptive and non-critical formulation: here is how an organization would design such a use directly with the framework, whatever its sector.

If an organization wished to govern this type of output directly with ACF, it could in particular:

- **document** a decisions register (ACF-08) linking each agent assertion to a real, verifiable source, checked before use;
- **declare** in a signed agentic constitution (ACF-03) the rule: never present an unverified source;
- **require** human supervision (ACF-05) checking the real existence of the sources before any binding filing;
- **map** the roles (ACF-01) to recall that the human who presents answers for it, whatever the delegated share;
- **rank** by criticality (ACF-02) the outputs according to the stake of the third party before whom they will be presented.

None of this is a technical reproach addressed to the agent: it is the way ACF would have made explicit, signable, and auditable the trace that a binding output must retain before being presented.

11. Similar cases (transferability)

This case is specific neither to law nor to the United States. It is transferable to any organization that uses an agent's output to **present it to a third party**:

-  law firms and legal professions
-  consulting, audit, and accounting
-  medical research and scientific writing
-  journalism and publishing
-  finance and due diligence
-  public administrations and regulatory reports

Everywhere, the same pattern: an agent that produces at scale, a traceability requirement back to a real source, and a human who remains accountable for what they present. This is what makes this document a **reference case**: it illustrates a concept, not just a case.

12. Confidence of the assessment

High. The case rests on a **published court decision** (sanction by Judge Castel, S.D.N.Y., June 22, 2023) and on legal analyses from reference firms. The facts are therefore established and verifiable by the reader. The ACF analysis draws on these public facts; it bears on governance, not on the internal technical details of the agent, which have no bearing on the reading proposed here.

Teacher guide

Pedagogical objective. Convey that the flaw is not the hallucination but the absence of a register and verification, and that delegating the research does not delegate accountability.

Session outline (90 min).

1. (15 min) Reading of the **two** cases in the kit: Mata v. Avianca (incident) and Perplexity (reference), without the analyses.
2. (20 min) In groups: define what a source register must guarantee (card ACF-08).
3. (20 min) Workshop: write the source-verification rule and the associated supervision (cards ACF-03 / 05).
4. (20 min) Mata / Perplexity comparison: assertion without a source versus verifiable citation.
5. (15 min) Synthesis: a binding agent output must trace back to a verifiable source.

Possible exam questions. See the "Governance questions" section.

Reminder: grading students' work is a pedagogical act of the teacher. This guide provides the reference analysis, not automatic grading.

Associated lab (ACF-LAB-028)

From this case, the student produces:

- the decision map (card ACF-01): agent, accountable party, non-delegable zone;
- the schema of a source register (card ACF-08): what to link, to which source, checked when;
- an agentic constitution (card ACF-03): the rule prohibiting the presentation of an unverified source;
- the supervision mechanism (card ACF-05): the human review verifying the real existence of the sources.

Appendix - Completed ACF cards

The official Toolkit cards used in this case, filled in with its public data. For illustration: an organization of the same profile can reuse and adapt them.



OBJECTIVE For the research delegated to ChatGPT, situate what the agent may suggest and what stays under the lawyer's accountability. Filled from public sources.

1 AGENTIC MATURITY SCALE

0 Classic automation Fixed rules, no autonomy.	1 Assisted agents The agent proposes, the human decides.	2 Governed agents The agent acts within a frame, under supervision. TARGET ~80%	3 Supervised autonomy The agent decides alone; control after the fact.
--	--	---	--

2 YOUR KEY DECISIONS (filled)

A Drafting assistance	B Research leads to verify	C Present a source to a court
STAKE Rephrase or structure a text already provided	STAKE Suggest case-law leads to check	STAKE File a citation as real and verified
IMPACT Reversible, no external assertion	IMPACT No value until verified	IMPACT Commits the lawyer and their accountability
DEADLINE Continuous	DEADLINE Continuous	DEADLINE -
AUTONOMY Autonomous	AUTONOMY Suggestive (to verify)	AUTONOMY NON-DELEGABLE

Reading: the agent could suggest leads; presenting a source as real before a court cannot be delegated. It is this line that was crossed.



OBJECTIVE

Define the agent's constitution: identity, standing principles and forbidden zones. Card filled with public case data.

1 AGENT IDENTITY & MISSION

AGENT NAME

ChatGPT, used for case-law research

ACCOUNTABLE (DDAO)

The lawyer and the firm

PRIMARY MISSION

Suggest references to support a brief, to be verified before use

SCOPE

Research leads; excluding the presentation of an unchecked source

2 CORE PRINCIPLES

- 1 Never present an unverified source as real
- 2 Attach every citation to a consultable decision before any filing
- 3 Have a human review before presenting to a third party
- 4 When in doubt about a source, abstain rather than assert

EXAMPLE PRINCIPLES

- Respect the legal and sector framework
- Protect data and confidentiality
- Stay transparent, every decision traceable
- When in doubt, escalate rather than decide

3 FORBIDDEN ZONES (ABSOLUTE)

- 1 File an unverified reference before a court
- 2 Provide fabricated excerpts to justify a citation
- 3 Present the agent's output as checked without it being so

EXAMPLE PROHIBITIONS

- X Commit spending above a threshold without validation
- X Share personal data with a third party
- X Change its own rules or scope on its own

Constitution filled retrospectively. Prohibitions 1 and 2 are exactly the ones whose absence led to the sanction of 22 June 2023 (Mata v. Avianca).



KILL SWITCH

Conditions for stopping and handing over to a human

LEVEL
USE
CASE
FILLED

Foundation
Control
ACF-CS-028
Mata v. Avianca

OBJECTIVE

Define the signals that block a filing or hand over to a human check. Card filled with the triggers relevant to the case.

1 TRIGGERS

Signal	Threshold	Action
Citation without a consultable source	Detected	Filing blocked until the source is verified
Source not findable by a third party	Reported	Reference removed, human review
Excerpt produced by the agent	Not backed by a real document	Use forbidden, escalation

2 PROCEDURE

- 1 Who decides the stop: the accountable lawyer (DDAO)
- 2 How: block the filing until every source is verified
- 3 Systematic check of the real existence of sources before presentation
- 4 Log every verification (ACF-08)

A block armed on any citation without a consultable source would have prevented the filing of the six non-existent decisions.

Card ACF-06 · Kill switch · completed with public case data.



DECISIONS REGISTER

Record who asserts what, to whom, on which source

LEVEL
USE
CASE
FILLED

Foundation
Traceability
ACF-CS-028
Mata v. Avianca

OBJECTIVE

Record every assertion taken from an agent and the real source it attaches to, before any engaging use. Card filled with the fields relevant to the case.

1 FIELDS TO RECORD

Field to record	Example, Mata v. Avianca case
Assertion produced	Six court decisions cited
Source invoked	No real, consultable source
Check performed	None before filing
Status	Non-existent references, fabricated excerpts
Accountable	The lawyer who presents to the court

2 USE OF THE REGISTER

- 1 Verify the real existence of a source before presenting it
- 2 Detect a citation not backed by a consultable document
- 3 Establish who accounts for the output: the lawyer, not the agent
- 4 Provide an enforceable trace in case of challenge

Without a register linking every citation to a real decision, there was nothing to trace, so nothing to defend before the judge.

Card ACF-08 · Decisions register · completed with public case data.

Disclaimer

This study is carried out **exclusively from public information** (published court decision and accessible legal analyses). It constitutes **neither an audit, nor a certification, nor an official evaluation** of the persons or entities cited. It applies the ACF framework for pedagogical and illustrative purposes. No internal, confidential, or non-public data was used. The trademarks cited belong to their respective owners.

Sources

- Decision of Judge P. Kevin Castel (U.S. District Court, S.D.N.Y.), *Mata v. Avianca, Inc.*, June 22, 2023
- Wikipedia, *Mata v. Avianca, Inc.* - https://en.wikipedia.org/wiki/Mata_v._Avianca,_Inc.
- Seyfarth Shaw, *Update on the ChatGPT Case: Counsel Who Submitted Fake Cases Are Sanctioned* (2023)

ACF Case Study ACF-CS-028 · v1 · working document, to be reviewed before any public distribution.