



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-029 · Case Study #029 · Édition 2026

Perplexity

How do you make an agent's output verifiable, so that it can be relied upon and answered for?

CARTE D'IDENTITÉ

SECTOR	COUNTRY	TECHNOLOGY	TYPE	AUTONOMY	CONFIDENCE
Search / AI	USA	Answer engine	Agentic AI	suggestive	Medium

OUTILS ACF ILLUSTRÉS DANS CE CAS

● ACF-01	● ACF-03	● ACF-05	● ACF-08	● ACF-00	● ACF-02
----------	----------	----------	----------	----------	----------

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

1. Executive summary

Perplexity is an AI-based answer engine: to a question, it responds in natural language, but by citing its sources. It is an agent that produces assertions meant to be used, like the delegated research in the Mata v. Avianca case, but with an opposite approach to traceability. Each statement carries a numbered citation that refers to a specific document, and the user can click to trace back to the original source.

Read through the lens of ACF, this case is a reference case on the **register and traceability**. The distinctive trait is that the citations are not added after the fact: the source is integrated into the construction of the answer, before generation itself. Trust does not rest on the idea that the agent never errs, but on the fact that each assertion can be verified and traced back to its origin. It is the exact counterpart of the Mata v. Avianca case (ACF-CS-028), where the output was linked to no real source.

2. Context

An agent that produces assertions meant to be used raises a simple governance question: how do you know where what it puts forward comes from? Without an answer, the output can be neither controlled nor defended the moment it binds. Perplexity illustrates one way of addressing this problem by design: making traceability not optional, but constitutive of the way the answer is built.

3. Public sources used

- Perplexity, "How does Perplexity work?" (help center).
- TrustRadius, "How Does Perplexity AI Work? Full Breakdown".
- ZipTie.dev, "How Perplexity AI Answers Work: Retrieval, Ranking, and Citation Pipeline".

4. Reconstructed AI architecture (from public sources)

Perplexity (answer engine with citations, as publicly documented).

- Nature: agent that answers in natural language from retrieved documents, and links its statements to sources.
- Function: produce an answer the user can reuse, exposing the origin of each element.
- Scope: the answer is meant to be verified, not taken at face value.

Traceability by design (what the facts describe):

- the **sources** (URL, metadata) are selected and **integrated before the generation** of the answer;
- each assertion carries a **numbered citation** referring to a specific document;
- the user can **click on the citation** and reach the original source to verify what is put forward;
- the trace is a **design property**, not an add-on tacked on after the fact.

5. ACF mapping

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

In ACF language, the architecture reads as follows:

ACF element	In the way Perplexity works
Named human (accountable)	The user who reuses the answer: it is they who verify via the citations and answer for the use they make of it.
Agent	The answer engine that retrieves, ranks, and writes with citations.
Platform	The underlying search and generation system.
Delegated decision	The production of a synthetic answer the user can reuse.
Autonomy level	Suggestive: the agent proposes a traced answer, the user verifies and decides on its use.
Non-delegable zones	The decision to use the information or not remains with the human, who controls it via the sources.
Rule / safeguard	Each assertion is linked to a clickable source, integrated from the very construction of the answer.
Control third party	The user themselves, to whom the trace gives the means to verify.
Traceability	High and by design: the numbered citation acts as a register, tracing back to a specific document.

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

6. Analysis

 **Tools used:** ACF-08 (register by design) · ACF-05 (facilitated supervision).

What the design does well (the reference design):

- **Systematic citation acts as a register.** Each assertion is linked to a verifiable source, integrated from the very construction of the answer. The output becomes controllable: one can trace back to the source, read it, contest it. This is exactly what was missing in the *Mata v. Avianca* case.
- **The trace lowers the cost of verification.** In one click, the user checks an assertion, spots a weak source, corrects it. This is what makes supervision realistic at scale: without a trace, verifying each statement would be impractical.
- **Tracing is not being right.** The important point, often misunderstood: the citation does not guarantee the truth of the source, it guarantees that it can be verified. This is traceability, not infallibility. The agent may cite a weak source, but one can see it, which was impossible in the *Mata* case.

What ACF names and generalizes:

- a **register by design**: the output is traceable back to its sources continuously, not after the fact;
- a **facilitated supervision**: the trace gives the means for human control, without replacing it;
- the transferable principle: traceability is a condition for relying on the output of a binding agent, not an add-on.

7. ACF toolkit used

This grid indicates the tools that the ACF analysis uses (✓) to read this reference 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	✓	Systematic citation as a register by design
ACF-05 · Supervision	✓	The trace lowers the cost of human control
ACF-01 · Decision map	✓	Locate the agent (suggestive) and the user (decides)
ACF-03 · Agentic constitution	→ SOON	Formalize the obligation to cite before asserting
ACF-02 · Criticality matrix	→ SOON	Calibrate the level of verification according to the stake

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

1. Is traceability an add-on, or a condition for relying on an agent's output?
2. Do our agents link each assertion to a verifiable source, from the very construction of the answer?
3. Does citing sources dispense with verifying, or does it only give the means to do so?
4. Does the trace lower the cost of control enough to make supervision realistic at scale?
5. When we reuse an agent's output, can we trace back to its source and answer for it?

9. General lessons

A trustworthy agent is not one that never errs, it is one whose every assertion can be verified. Perplexity shows that traceability can be a design property: the source is there before the answer, and the user can trace back to the origin. This trace does not replace human control, it makes it possible by lowering its cost.

The transferable lesson: traceability does not dispense with verifying, it gives the means to do so. A binding agent output must trace back to its source.

10. What ACF would do (reference design)

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

If an organization wished to design this type of assertion-producing agent directly with ACF, it could in particular:

- **document** a register by design (ACF-08) linking each assertion to a verifiable source, integrated from the very construction of the answer;
- **link** this register to human supervision (ACF-05) that leverages the trace to control at low cost;
- **map** the roles (ACF-01) to recall that the user who reuses the output answers for it;
- **declare** in an agentic constitution (ACF-03) the obligation to cite before asserting;
- **calibrate** by criticality (ACF-02) the level of verification required according to the stake of the use.

None of this is a commendation addressed to Perplexity: it is the way ACF would make explicit, signable, and auditable the traceability that an assertion-producing agent must maintain for it to be relied upon.

11. Similar cases (transferability)

This case is specific neither to Perplexity nor to answer engines. It is transferable to any organization that deploys an agent **producing assertions to be reused**:

-  answer engines and research assistants
-  enterprise document synthesis tools
-  decision support (legal, medical, financial)
-  market watch and competitive intelligence
-  assisted writing and technical documentation
-  customer support based on a knowledge base

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

12. Confidence of the assessment

Medium. The case rests on public sources (Perplexity's help center and third-party technical analyses) that document the principle of citation by design. This governance principle is therefore established and verifiable in its broad outlines. On the other hand, the internal detail of the retrieval and ranking pipeline is not fully public and may evolve. The analysis bears on the traceability logic displayed, with these details having no bearing on the reading proposed here.

Teacher guide

Pedagogical objective. Convey traceability as a condition of trust: citing allows verification, without guaranteeing the truth or replacing human control.

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: describe systematic citation as a register (card ACF-08).
3. (20 min) Debate: "Does citing sources dispense with verifying?" (expected answer: no, it gives the means to do so).
4. (20 min) Mata / Perplexity comparison: without a trace versus traceable by design.
5. (15 min) Synthesis: a binding agent output must be traceable back to its 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-029)

From this case, the student produces:

- the decision map (card ACF-01): suggestive agent, user who decides;
 - the schema of a register by design (card ACF-08): what the numbered citation links and enables;
 - the supervision mechanism (card ACF-05): what the trace makes verifiable in one click;
 - a written comparison with the Mata v. Avianca case: what changes on the side of source traceability.
-

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

Situate the engine's role (offer a traced answer) and the user's role (verify and decide on the use). 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

Traced answer to a question

STAKE

Answer in natural language with citations

IMPACT

Reversible, the user verifies

DEADLINE

Continuous

AUTONOMY

Suggestive

B

Referral to the source

STAKE

Attach every statement to a clickable document

IMPACT

Provides the means of control

DEADLINE

Continuous

AUTONOMY

Suggestive (cites the source)

C

Decide to use the information

STAKE

Take up and commit on the answer

IMPACT

Stays with the human who accounts for it

DEADLINE

-

AUTONOMY

NON-DELEGABLE

Reading: the agent is **suggestive**, it offers a verifiable answer; the decision to use it stays with the user, who controls through the citations.



OBJECTIVE

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

1 AGENT IDENTITY & MISSION

AGENT NAME

Perplexity, answer engine with citations

ACCOUNTABLE (DDAO)

The user who takes up the answer

PRIMARY MISSION

Produce a synthetic answer, traceable back to its sources

SCOPE

Answer meant to be verified; excluding the use decision, which stays human

2 CORE PRINCIPLES

- 1 Cite before asserting: integrate the source from the construction of the answer
- 2 Attach every statement to a precise, clickable document
- 3 Make the trace verifiable by the user in one click
- 4 Do not present the citation as a guarantee of truth

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 Assert without attaching the statement to a source
- 2 Present a trace as proof of the source's accuracy
- 3 Decide, in the user's place, on the use of the information

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 read from public sources. The distinctive trait: the citation is not added afterward, it is built in before generation.



OBJECTIVE

Link traceability by design to a realistic human control at scale. Card filled with the publicly documented elements of the case.

1 SUPERVISION INDICATORS

Indicator	Value / target (public sources)
Traceability of statements	Numbered citation per assertion
Origin of sources	Selected before generation
Verifiability	Clickable source, back to the document
Cost of control	Lowered: verification in one click
Guarantee of truth	No: the trace lets you verify, not be right

2 DEVICE

- 1 Traceability by design: the source is there before the answer
- 2 Human control: the user goes back to the source and judges its quality
- 3 Supervision owner: the user who takes up the output
- 4 Calibrate the level of verification to the stakes of the use (ACF-02)

The trace does not replace human control: it makes it possible by lowering its cost. Without it, verifying every statement at scale would be impractical.



OBJECTIVE

Show how systematic citation acts as a register: every assertion goes back to a verifiable source. Card filled with the public data of the case.

1 FIELDS TO RECORD

Field to record	In Perplexity's operation
Assertion produced	Every statement of the answer
Attached source	Precise document, via a numbered citation
Moment of attachment	Before generation, by design
Verification	Click on the citation, back to the source
Use owner	The user who takes up the answer

2 USE OF THE REGISTER

- 1 Go back from every assertion to the source that grounds it
- 2 Spot a weak source and correct before taking it up
- 3 Make supervision realistic at scale, at low cost
- 4 Establish who accounts for the use: the user, not the agent

The numbered citation acts as a **register by design**: exactly what was missing in the Mata v. Avianca case (ACF-CS-028).

Disclaimer

This study is carried out **exclusively from public information**. It constitutes **neither an audit, nor a certification, nor an official evaluation** of Perplexity. 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

- Perplexity, *How does Perplexity work?* (help center) - <https://www.perplexity.ai/hub/faq/how-does-perplexity-work>
- TrustRadius, *How Does Perplexity AI Work? Full Breakdown*
- ZipTie.dev, *How Perplexity AI Answers Work: Retrieval, Ranking, and Citation Pipeline*

ACF Case Study ACF-CS-029 · v1 · working document, to be reviewed before any public distribution.