



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-003 · Case Study #003 · Édition 2026

Air Canada

When a conversational agent gets it wrong, who answers for it?

CARTE D'IDENTITÉ

SECTOR Air transport	COUNTRY Canada	TECHNOLOGY Conversational agent	TYPE Public chatbot
AUTONOMY autonomous (de facto)	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

Air Canada, the country's main airline, offers on its website a conversational agent to answer customers' routine questions. In November 2022, this chatbot asserts to a passenger a **false** bereavement fare policy, contrary to the official page of the same site. The passenger relies on it, buys a full fare, then is denied the refund. The tribunal rules against Air Canada (Moffatt v. Air Canada, 2024 BCCRT 149).

Read through the lens of ACF, this case is the **"incident"** counterpart of Morgan Stanley: both deal with the same object, the **non-delegation of accountability**, by opposite paths. Where Morgan Stanley holds its non-delegable zones by keeping autonomy low, Air Canada let an agent **bind the company** on information that should never have been delegable. The value of ACF here is **corrective**: it names, from the design stage, the mapping error that cost the lawsuit.

2. Context

Like most large service companies, Air Canada deploys a chatbot to handle simple requests at scale (baggage, fares, policies, bookings) without mobilizing a human advisor each time. The intent is legitimate and common. The question the case raises is just as much so: when this agent asserts something binding and gets it wrong, does the company answer for it?

3. Public sources used

- Moffatt v. Air Canada, 2024 BCCRT 149 (British Columbia Civil Resolution Tribunal), decision published February 2024.
- McCarthy Tétrault, "Moffatt v. Air Canada: A Misrepresentation by an AI Chatbot".
- American Bar Association, "BC Tribunal Confirms Companies Remain Liable for Information Provided by AI Chatbot".
- Dentons Data, "Airline ordered to compensate a B.C. man because its chatbot provided inaccurate information".

4. Reconstructed AI architecture (from public sources)

Air Canada customer service chatbot (in service at the time of the facts, November 2022).

- Nature: conversational agent integrated into Air Canada's public website, freely accessible to customers.
- Function: answer routine questions (fares, policies, bookings). In practice, it **asserted** policies, without limiting itself to directing to the official source.
- Actual scope observed: the chatbot produced a **binding** answer (a refund policy) that was **unsourced** and **contrary** to the site's official page.

The breaking point. On November 11, 2022, Jake Moffatt, booking urgently following a death, asks the chatbot about bereavement fares. The chatbot replies that he can request the reduced fare **after the flight, within 90 days**. The real policy grants **no** retroactive refund: the discount must be approved **before** travel. Moffatt buys the full fare relying on the chatbot, then claims the difference.

Missing safeguards (what the facts reveal):

- no constraint preventing the agent from **asserting an unverified policy**;
- no **consistency** rule with the site's official pages;
- no **escalation** or blocking on binding information;
- no usable **trace** of the answer given, the customer concerned, and the source invoked.

5. ACF mapping

Tools used: ACF-01 (decision map) · ACF-03 (agentic constitution) · ACF-06 (kill switch) · ACF-08 (register).

In ACF language, the incident reads as follows:

ACF element	In the Air Canada deployment
Named human (accountable)	Air Canada, the company. It is the one that answers for everything its agent says (tribunal: the chatbot is part of the website).
Agent	The customer service chatbot of the public site.
Platform	The underlying conversational system integrated into the site.
Delegated decision	In effect, the agent decided and bound the company on a refund policy, a decision that should not have been delegated to it.
Autonomy level	Autonomous on binding information: the chatbot asserted alone, without review or source.
Non-delegable zones	Asserting a fare policy or binding the company on a refund. This is precisely the boundary that was crossed.
Rule / safeguard	No enforceable safeguard: neither a source constraint, nor consistency with the official page, nor escalation.
Control third party	After the fact, the tribunal (British Columbia Civil Resolution Tribunal).
Traceability	No decisions register in the ACF sense: the faulty answer was neither timestamped, nor linked to a source, nor reconstructable ahead of the dispute.

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

6. Analysis

Tools used: ACF-01 (mapping) · ACF-03 (non-delegable zones) · ACF-06 (stop) · ACF-08 (register).

What the incident reveals (the design error):

- **A non-delegable zone crossed.** Informing about a stable fact is delegable; **asserting a binding policy** is not. The deployment did not draw this boundary, and the agent bound the company where a named human should have stayed in the loop.
- **Accountability was not diluted, it was engaged.** Air Canada's argument, "the chatbot is a separate legal entity, responsible for its own acts", was rejected. An agent never dilutes the accountability of the one who deploys it: it **engages** it. This is the ACF formula: you delegate execution, never accountability.
- **The absence of a trace worsened the case.** Without a register linking the answer to a source and a customer, the error was neither detectable nor correctable at scale before the dispute.

What ACF would have made explicit (and enforceable):

- an **agentic constitution** prohibiting asserting an unsourced policy, contradicting an official page, and committing a refund without validation;
- a **stop trigger**: as soon as the agent formulates a binding unsourced policy, the answer is blocked and escalated to a human;
- a **decisions register** timestamping which answer was given, to which customer, on which source.

7. ACF toolkit used

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

Tool	Status	Role in this case
ACF-01 · Decision map	✓	Situate the agent, the accountable party, and the decision engaged
ACF-03 · Agentic constitution	✓	Set the forbidden zones (unsourced policy, contradiction, commitment)
ACF-06 · Kill switch	✓	Block and escalate any binding unsourced assertion
ACF-08 · Decisions register	✓	Timestamp the answer, the customer, the source, to correct at scale
ACF-02 · Criticality matrix	→ SOON	Classify topics by criticality (informative vs binding)
ACF-05 · Supervision	→ SOON	Link the monitoring of answers to continuous follow-up

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

1. Can a company hide behind "the chatbot said it"? The tribunal answered; and us?
2. Which information, in our agent, **binds** the company, and which remains purely informative?
3. Where have we placed, explicitly, the boundary between informing and deciding?
4. Can our agent contradict an official page without any rule preventing it?
5. Do we have a trace, reconstructable and enforceable, of who answered what and on which source?

9. General lessons

A conversational agent does not fail because it "speaks badly": it fails when it is allowed to **decide** something that should have remained non-delegable. Air Canada did not lack technology, it lacked **mapping**: no one had declared, upstream, that asserting a binding policy was not the agent's remit.

The transferable lesson: delegating the execution of a conversation never delegates the accountability for what is asserted. Governance is cheaper than the incident.

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 deploy this type of agent directly with ACF, it could in particular:

- **map** each type of answer on the autonomy scale (ACF-01) and distinguish the informative from the binding;
- **declare** in a signed agentic constitution (ACF-03) the non-delegable zones: never assert an unsourced policy, never contradict an official page, never commit a refund without validation;
- **define** a stop mechanism (ACF-06) and its trigger: any binding unsourced assertion is blocked and escalated;
- **document** a decisions register (ACF-08) linking each answer to its source and to the customer concerned;
- **classify** topics by criticality (ACF-02) to calibrate, topic by topic, what the agent can say alone.

None of this is a technical reproach: it is the way ACF would have made **explicit, signable, and auditable** the boundary that a prudent deployment must hold before opening an agent to the public.

11. Similar cases (transferability)

This case is specific neither to Air Canada nor to air transport. It is transferable to any organization that exposes to the public a **conversational agent capable of asserting binding policies**:

-  airlines and carriers
-  telecoms and energy providers
-  banks and insurers (mass customer service)
-  e-commerce and retail
-  public services and administrations
-  healthcare (appointment booking, patient information)

Everywhere, the same pattern: an agent that informs at scale, a boundary (the policy, the price, the commitment) never to be crossed without a named human, and a trace to keep. This is what makes this document a **reference case**: it illustrates a concept, not just a company.

12. Confidence of the assessment

High. The case rests on a **published court decision** (Moffatt v. Air Canada, 2024 BCCRT 149) 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 concerns **governance**, not the agent's internal technical details, which are non-public and without bearing on the reading proposed.

Teacher guide

Pedagogical objective. Make clear that an agent does not dilute accountability but engages it, and where to place the boundary between informing and deciding.

Session outline (90 min).

1. (15 min) Reading of the **two** cases of the kit: Morgan Stanley (reference) and Air Canada (incident), without the analyses.
2. (20 min) In groups: map Air Canada in ACF language (who executes, who answers, which non-delegable zone).
3. (20 min) Debate: "Can a company hide behind its chatbot?" (expected answer: no, backed by the tribunal's decision).
4. (20 min) Projection: write the three constitution rules that would have avoided the incident.
5. (15 min) Synthesis: is accountability delegable? No, you delegate execution, not accountability.

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-003)

From this case, the student produces:

- the decision map of the chatbot (card ACF-01): agent, accountable party, non-delegable zone;
 - an agentic constitution (card ACF-03): three enforceable prohibitions;
 - the definition of a kill switch (card ACF-06) and its trigger;
 - the schema of a decisions register (card ACF-08): what to timestamp and why.
-

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

Map the chatbot's replies and set, for each, the autonomy level based on what it commits, not on how frequent it is. 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 Stable factual information

STAKE

Answer on schedules, baggage, procedures

IMPACT

Reversible, no commitment

DEADLINE

Continuous

AUTONOMY

Autonomous

B Engaging topic, orientation

STAKE

Point to the official page on a fare or policy

IMPACT

No assertion of its own

DEADLINE

Continuous

AUTONOMY

Suggestive (cites the source)

C Assert a policy, commit a refund

STAKE

Decide a fare, promise a refund

IMPACT

Commits the company and its accountability

DEADLINE

-

AUTONOMY

NON-DELEGABLE

Reading: it is this setting, made **before** deployment, that would have prevented the agent from asserting a refund policy on its own.



OBJECTIVE

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

1 AGENT IDENTITY & MISSION

AGENT NAME

Air Canada customer service chatbot

ACCOUNTABLE (DDAO)

To be designated

PRIMARY MISSION

Answer customers' routine questions, at scale

SCOPE

Factual information; excluding policy assertions and financial commitments

2 CORE PRINCIPLES

- 1 Never assert an unsourced policy
- 2 Always point to the official page on an engaging topic
- 3 When in doubt, escalate rather than decide
- 4 Signal to the customer that they are talking to an AI

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 Invent or assert an unverified fare policy
- 2 Contradict an official page of the website
- 3 Commit the company to a refund without human validation

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 precisely the ones whose absence led to the ruling (Moffatt v. Air Canada, 2024).



KILL SWITCH

Conditions for stopping and handing over to a human

LEVEL
USE
CASE
FILLED

Foundation
Control
ACF-CS-003
Air Canada

OBJECTIVE

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

1 TRIGGERS

Signal	Threshold	Action
Reply asserting an engaging policy	Unsourced, outside the validated base	Reply blocked, human escalation
Inconsistency with an official page	Detected	Topic suspended, review
Complaint rate on a topic	Above threshold	Topic removed from the agent's scope

2 PROCEDURE

- 1 Who decides the stop: the DDAO
- 2 How: block the engaging reply, holding message, human escalation
- 3 Regular testing of the device on sensitive topics
- 4 Log every block (ACF-08)

It is precisely this block, armed on any assertion of an unsourced policy, that would have avoided the false information at the origin of the dispute.

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



OBJECTIVE

Record every engaging reply to be able to reconstruct it, correct it at scale and account for it. Card filled with the fields relevant to the case.

1 FIELDS TO RECORD

Field to record	Example, Air Canada case
Timestamp	11 November 2022
Customer / session	Query about the bereavement fare
Reply given	Refund possible within 90 days after the flight
Source invoked	None (unsourced reply)
Status	Contrary to the website's official page

2 USE OF THE REGISTER

- 1 Detect an unsourced reply before it repeats
- 2 Correct at scale all similar cases
- 3 Establish who accounts for the reply: the company
- 4 Provide enforceable evidence in case of dispute

Without a register, the error was neither detectable nor correctable before the trial. The register turns an isolated incident into a systemic fix.

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 Air Canada. 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

- Moffatt v. Air Canada, 2024 BCCRT 149 (British Columbia Civil Resolution Tribunal) - <https://www.canlii.org/en/bc/bccrt/doc/2024/2024bccrt149/2024bccrt149.html>
- McCarthy Tétrault, *Moffatt v. Air Canada: A Misrepresentation by an AI Chatbot*
- American Bar Association, *BC Tribunal Confirms Companies Remain Liable for Information Provided by AI Chatbot*
- Dentons Data, *Airline ordered to compensate a B.C. man because its chatbot provided inaccurate information*

ACF Case Study ACF-CS-003 · v1 · working document, to be reviewed before any public distribution.