



ACF EN 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 (l'avocat et ChatGPT)

Peut-on présenter la production d'un agent sans trace vérifiable de ses sources ?

CARTE D'IDENTITÉ

SECTEUR Juridique	PAYS USA	TECHNOLOGIE Assistant de recherche	TYPE IA générative	AUTONOMIE autonome
CONFIANCE Moyenne				

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. Résumé exécutif

En 2023, dans une affaire civile aux États-Unis (Mata v. Avianca), des avocats préparent un mémoire juridique et délèguent à ChatGPT la recherche de jurisprudence. L'agent fabrique six décisions de justice entièrement fausses, avec noms d'affaires, citations et extraits inventés. Les avocats ne vérifient pas et déposent ces références au tribunal. Sommés de produire les textes, ils fournissent des extraits eux-mêmes fabriqués. Le 22 juin 2023, le juge les sanctionne d'une amende de 5 000 dollars.

Lu au prisme de l'ACF, ce cas est le pendant « **incident** » de Perplexity (ACF-CS-029) : les deux traitent le même objet, le **registre et la traçabilité**, par des chemins opposés. Le problème n'est pas seulement que l'agent ait halluciné, l'hallucination est attendue. La faute de gouvernance est d'avoir présenté sa production **sans registre ni vérification des sources**. Une affirmation d'agent qui engage n'a de valeur que si elle est traçable jusqu'à une source réelle.

2. Contexte

Déléguer une recherche documentaire à un agent conversationnel est une pratique courante et légitime. La difficulté surgit quand cette production est présentée à un tiers qui l'attend vérifiable. Une citation de jurisprudence n'a de sens que si elle renvoie à une décision réelle, consultable. Le cas Mata v. Avianca montre ce qui se passe quand cette exigence de traçabilité disparaît entre la production de l'agent et le dépôt devant le tribunal.

3. Sources publiques utilisées

- Décision du juge P. Kevin Castel (U.S. District Court, S.D.N.Y.), 22 juin 2023.
- Wikipedia, « Mata v. Avianca, Inc. » (faits et décision sourcés).
- Seyfarth Shaw, « Update on the ChatGPT Case: Counsel Who Submitted Fake Cases Are Sanctioned » (2023).

4. Architecture IA reconstituée (à partir de sources publiques)

Recherche de jurisprudence déléguée à ChatGPT (affaire Mata v. Avianca, 2023).

- Nature : agent conversationnel généraliste utilisé comme outil de recherche juridique.
- Fonction attendue : produire des références de jurisprudence à l'appui d'un mémoire.
- Périmètre réel constaté : l'agent a produit six décisions **inexistantes**, avec des citations et des extraits **inventés**, présentés comme réels.

Le point de rupture. Les avocats intègrent ces références au mémoire déposé, sans les vérifier. Quand la partie adverse signale ne retrouver aucune des décisions citées, le tribunal réclame les textes. Au lieu de reconnaître l'erreur, un avocat fournit des extraits, eux aussi fabriqués par l'agent. À aucun moment la chaîne d'affirmations n'est adossée à une source réelle et consultable.

Garde-fous absents (ce que les faits révèlent) :

- aucune **vérification** de l'existence réelle des sources avant dépôt ;
- aucune **règle** interdisant de présenter une référence non contrôlée ;
- aucun **registre** reliant chaque citation à une décision réelle et consultable ;
- aucune **relecture humaine** exigée avant de présenter la production à un tiers.

5. Cartographie ACF

 **Outils mobilisés** : ACF-01 (cartographie) · ACF-03 (constitution agentique) · ACF-05 (supervision) · ACF-08 (registre).

En langage ACF, l'incident se lit ainsi :

Élément ACF	Dans le cas Mata v. Avianca
Humain nommé (responsable)	L'avocat et le cabinet : c'est lui qui répond de ce qu'il présente au tribunal, sources comprises.
Agent	ChatGPT, utilisé pour produire des références juridiques.
Plateforme	Le modèle conversationnel généraliste sous-jacent.
Décision déléguée	La recherche de jurisprudence destinée à un dépôt judiciaire. La délégation de la recherche ne délègue pas la responsabilité.
Niveau d'autonomie	Autonome de fait : la production a été reprise sans contrôle, comme si elle était vérifiée.
Zones non déléguables	Garantir qu'une source présentée à un tribunal existe et est vérifiable. C'est la frontière qui a été franchie.
Règle / garde-fou	Aucune : ni vérification des sources, ni interdiction de présenter une référence non contrôlée.
Tiers de contrôle	A posteriori, le tribunal (juge Castel, S.D.N.Y.).
Traçabilité	Nulle : aucune citation ne renvoyait à une décision réelle. Il n'y avait rien à tracer, donc rien à défendre.

6. Analyse

 **Outils mobilisés** : ACF-08 (registre) · ACF-03 (règle de vérification) · ACF-05 (supervision) · ACF-01 (responsabilité).

Ce que l'incident révèle (l'erreur de gouvernance) :

- **Le vrai défaut n'est pas l'hallucination, c'est l'absence de trace.** Un agent généraliste peut inventer des références : c'est un risque connu. La faute est d'avoir présenté cette production à un tribunal **sans registre reliant chaque citation à une source réelle**, et sans l'avoir contrôlée avant usage.
- **Déléguer la recherche ne délègue pas la responsabilité.** L'avocat répond de ce qu'il dépose, sources comprises. L'agent produit ; l'humain qui présente sans vérifier répond. C'est la non-délégation de la responsabilité.
- **Une affirmation sans source vérifiable ne vaut rien dès qu'elle engage.** Sans trace remontant à une décision consultable, la production n'était ni contrôlable, ni corrigeable, ni défendable. Fournir ensuite des extraits fabriqués n'a fait qu'aggraver le défaut initial.

Ce que l'ACF aurait rendu explicite (et opposable) :

- un **registre** garantissant que chaque référence est rattachée à une source réelle, vérifiable et contrôlée avant tout usage ;
- une **règle de constitution** interdisant de présenter une source non vérifiée ;
- une **supervision** imposant une relecture humaine qui vérifie l'existence réelle des sources avant dépôt.

7. Toolkit ACF mobilisé

Cette grille indique les outils que l'analyse ACF mobilise (✓) pour lire et corriger ce cas, et ceux à prévoir (→
SOON). Elle ne signifie pas que l'entreprise utilise l'ACF : elle montre quels instruments le référentiel apporterait.

Outil	Statut	Rôle dans ce cas
ACF-08 · Registre des décisions	✓	Rattacher chaque référence à une source réelle et vérifiable
ACF-03 · Constitution agentique	✓	Poser la règle : ne jamais présenter une source non vérifiée
ACF-05 · Supervision	✓	Imposer une relecture humaine contrôlant l'existence des sources
ACF-01 · Carte décisionnelle	✓	Situer le responsable : l'avocat, pas l'agent
ACF-02 · Matrice de criticité	→ SOON	Classer par enjeu ce qui exige vérification avant usage
ACF-06 · Kill switch	→ SOON	Bloquer un dépôt tant que les sources ne sont pas vérifiées

8. Questions de gouvernance (ce qu'un comité IA devrait se poser)

1. Une affirmation d'agent sans source vérifiable a-t-elle une valeur dès qu'elle engage devant un tiers ?
2. À qui incombe la responsabilité : à l'agent qui hallucine, ou à celui qui présente sans vérifier ?
3. Avons-nous un registre reliant chaque production d'agent à une source réelle et consultable ?
4. Une relecture humaine vérifie-t-elle l'existence réelle des sources avant tout usage engageant ?
5. Nos règles interdisent-elles, explicitement, de présenter une source non contrôlée ?

9. Enseignements généraux

Un agent n'a pas échoué parce qu'il a halluciné : il hallucine par nature. Le défaut de gouvernance est d'avoir repris sa production sans trace ni vérification, puis de l'avoir présentée à un tribunal. Déléguer la recherche ne décharge personne de la responsabilité de ce qui est déposé. Ce qui manquait n'était pas un meilleur prompt, mais un registre et un contrôle systématique des sources.

La leçon transférable : une production d'agent qui engage ne vaut que si elle remonte à une source vérifiable. Le registre est moins cher que la sanction.

10. Ce que ferait l'ACF (conception de référence)

Formulation prescriptive et non critique : voici comment une organisation concevrait un tel usage directement avec le référentiel, quel que soit son secteur.

Si une organisation souhaitait encadrer ce type de production directement avec l'ACF, elle pourrait notamment :

- **documenter** un registre des décisions (ACF-08) reliant chaque affirmation d'agent à une source réelle, vérifiable et contrôlée avant usage ;
- **déclarer** dans une constitution agentique signée (ACF-03) la règle : ne jamais présenter une source non vérifiée ;
- **imposer** une supervision humaine (ACF-05) contrôlant l'existence réelle des sources avant tout dépôt engageant ;
- **cartographier** les rôles (ACF-01) pour rappeler que l'humain qui présente répond, quelle que soit la part déléguée ;
- **classer** par criticité (ACF-02) les productions selon l'enjeu du tiers devant qui elles seront présentées.

Rien de tout cela n'est un reproche technique adressé à l'agent : c'est la manière dont l'ACF aurait rendu explicite, signable et auditable la trace qu'une production engageante doit conserver avant d'être présentée.

11. Cas similaires (transférabilité)

Ce cas n'est propre ni au droit ni aux États-Unis. Il est transférable à toute organisation qui reprend la production d'un agent pour la **présenter à un tiers** :

-  cabinets d'avocats et professions juridiques
-  conseil, audit et expertise comptable
-  recherche médicale et rédaction scientifique
-  journalisme et édition
-  finance et due diligence
-  administrations et rapports réglementaires

Partout, le même schéma : un agent qui produit à grande échelle, une exigence de traçabilité jusqu'à une source réelle, et un humain qui reste responsable de ce qu'il présente. C'est ce qui fait de ce document un **cas de référence** : il illustre un concept, pas seulement une affaire.

12. Confiance de l'évaluation

Élevée. Le cas repose sur une **décision de justice publiée** (sanction du juge Castel, S.D.N.Y., 22 juin 2023) et sur des analyses juridiques de cabinets de référence. Les faits sont donc établis et vérifiables par le lecteur. L'analyse ACF s'appuie sur ces faits publics ; elle porte sur la gouvernance, pas sur les détails techniques internes de l'agent, sans incidence sur la lecture proposée.

Guide enseignant

Objectif pédagogique. Faire comprendre que le défaut n'est pas l'hallucination mais l'absence de registre et de vérification, et que déléguer la recherche ne délègue pas la responsabilité.

Déroulé de séance (90 min).

1. (15 min) Lecture des **deux** cas du kit : Mata v. Avianca (incident) et Perplexity (référence), sans les analyses.
2. (20 min) En groupes : définir ce qu'un registre de sources doit garantir (fiche ACF-08).
3. (20 min) Atelier : écrire la règle de vérification des sources et la supervision associée (fiches ACF-03 / 05).
4. (20 min) Comparaison Mata / Perplexity : affirmation sans source contre citation vérifiable.
5. (15 min) Synthèse : une production d'agent qui engage doit remonter à une source vérifiable.

Questions d'examen possibles. Voir la section « Questions de gouvernance ».

Rappel : la correction du travail des étudiants est un acte pédagogique du professeur. Le présent guide fournit l'analyse de référence, pas une notation automatique.

Lab associé (ACF-LAB-028)

À partir de ce cas, l'étudiant produit :

- la cartographie décisionnelle (fiche ACF-01) : agent, responsable, zone non délégable ;
- le schéma d'un registre des sources (fiche ACF-08) : quoi rattacher, à quelle source, contrôlé quand ;
- une Constitution agentique (fiche ACF-03) : la règle interdisant de présenter une source non vérifiée ;
- le dispositif de supervision (fiche ACF-05) : la relecture humaine vérifiant l'existence réelle des sources.

Annexe - Fiches ACF complétées

Les fiches officielles du Toolkit mobilisées dans ce cas, remplies avec ses données publiques. À titre illustratif : une organisation du même profil peut les reprendre et les adapter.



OBJECTIF

Situer, pour la recherche déléguée à ChatGPT, ce que l'agent peut proposer et ce qui reste sous la responsabilité de l'avocat. Rempli à partir des sources publiques.

1 ÉCHELLE DE MATURITÉ AGENTIQUE

0

Automatisation classique

Règles fixes, aucune autonomie.

1

Agents assistés

L'agent propose, l'humain décide.

2

Agents gouvernés

L'agent agit dans un cadre, sous supervision.

CIBLE ~80 %

3

Autonomie supervisée

L'agent décide seul ; contrôle a posteriori.

2 VOS DÉCISIONS CLÉS (rempli)

A

Aide à la rédaction

ENJEU

Reformuler ou structurer un texte déjà fourni

IMPACT

Réversible, sans affirmation externe

ÉCHÉANCE

Continu

AUTONOMIE

Autonome

B

Recherche documentaire à vérifier

ENJEU

Proposer des pistes de jurisprudence à contrôler

IMPACT

Aucune valeur tant que non vérifiée

ÉCHÉANCE

Continu

AUTONOMIE

Suggestif (à vérifier)

C

Présenter une source à un tribunal

ENJEU

Déposer une citation comme réelle et vérifiée

IMPACT

Engage l'avocat et sa responsabilité

ÉCHÉANCE

-

AUTONOMIE

NON DÉLÉGABLE

Lecture : l'agent pouvait proposer des pistes ; présenter une source comme réelle devant un tribunal ne se délègue pas. C'est cette ligne qui a été franchie.



OBJECTIF

Définir la constitution de l'agent : identité, principes permanents et zones absolument interdites. Fiche remplie avec les données publiques du cas.

1 IDENTITÉ & MISSION DE L'AGENT

NOM DE L'AGENT

ChatGPT, utilisé en recherche de jurisprudence

RESPONSABLE (DDAO)

L'avocat et le cabinet

MISSION PRINCIPALE

Proposer des références à l'appui d'un mémoire, à vérifier avant usage

PÉRIMÈTRE

Pistes de recherche ; hors présentation d'une source non contrôlée

2 PRINCIPES FONDAMENTAUX

- 1 Ne jamais présenter une source non vérifiée comme réelle
- 2 Rattacher chaque citation à une décision consultable avant tout dépôt
- 3 Faire relire par un humain avant présentation à un tiers
- 4 En cas de doute sur une source, s'abstenir plutôt que d'affirmer

EXEMPLES DE PRINCIPES

- Respecter le cadre légal et sectoriel
- Protéger les données et la confidentialité
- Rester transparent, toute décision traçable
- En cas de doute, escalader plutôt que décider

3 ZONES INTERDITES (ABSOLUES)

- 1 Déposer une référence non vérifiée devant un tribunal
- 2 Fournir des extraits fabriqués pour justifier une citation
- 3 Présenter la production de l'agent comme contrôlée sans l'avoir été

EXEMPLES D'INTERDITS

- X Engager une dépense au-delà d'un seuil sans validation
- X Communiquer des données personnelles à un tiers
- X Modifier seul ses propres règles ou son périmètre

Constitution remplie rétrospectivement. Les interdictions 1 et 2 sont exactement celles dont l'absence a conduit à la sanction du 22 juin 2023 (Mata v. Avianca).



OBJECTIF

Définir les signaux qui bloquent un dépôt ou basculent vers un contrôle humain. Fiche remplie avec les déclencheurs pertinents pour le cas.

1 DÉCLENCHEURS

Signal	Seuil	Action
Citation sans source consultable	Détectée	Dépôt bloqué tant que la source n'est pas vérifiée
Source introuvable par un tiers	Signalée	Retrait de la référence, revue humaine
Extrait produit par l'agent	Non adossé à un document réel	Interdiction d'usage, escalade

2 PROCÉDURE

- 1 Qui décide l'arrêt : l'avocat responsable (DDAO)
- 2 Comment : blocage du dépôt tant que chaque source n'est pas vérifiée
- 3 Contrôle systématique de l'existence réelle des sources avant présentation
- 4 Journalisation de chaque vérification (ACF-08)

Un blocage armé sur toute citation sans source consultable aurait empêché le dépôt des six décisions inexistantes.



OBJECTIF

Consigner chaque affirmation reprise d'un agent et la source réelle à laquelle elle se rattache, avant tout usage engageant. Fiche remplie avec les champs pertinents pour le cas.

1 CHAMPS À TRACER

Champ à tracer	Exemple, cas Mata v. Avianca
Affirmation produite	Six décisions de justice citées
Source invoquée	Aucune source réelle et consultable
Contrôle effectué	Aucun avant dépôt
Statut	Références inexistantes, extraits fabriqués
Responsable	L'avocat qui présente au tribunal

2 USAGE DU REGISTRE

- 1 Vérifier l'existence réelle d'une source avant de la présenter
- 2 Détecter une citation non adossée à un document consultable
- 3 Établir qui répond de la production : l'avocat, pas l'agent
- 4 Fournir une trace opposable en cas de contestation

Sans registre reliant chaque citation à une décision réelle, il n'y avait rien à tracer, donc rien à défendre devant le juge.

Disclaimer

Cette étude est réalisée **exclusivement à partir d'informations publiques** (décision de justice publiée et analyses juridiques accessibles). Elle ne constitue **ni un audit, ni une certification, ni une évaluation officielle** des personnes ou entités citées. Elle applique le référentiel ACF à des fins pédagogiques et illustratives. Aucune donnée interne, confidentielle ou non publique n'a été utilisée. Les marques citées appartiennent à leurs détenteurs respectifs.

Sources

- Décision du juge P. Kevin Castel (U.S. District Court, S.D.N.Y.), *Mata v. Avianca, Inc.*, 22 juin 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 · document de travail, à relire avant toute diffusion publique.