



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-005 · Case Study #005 · Édition 2026

Waymo (Alphabet)

When an autonomous agent gets it wrong, can a company stop itself before being forced to?

CARTE D'IDENTITÉ

SECTOR Mobility	COUNTRY USA	TECHNOLOGY Autonomous robotaxi	TYPE Agentic AI	AUTONOMY supervised autonomous
CONFIDENCE High				

OUTILS ACF ILLUSTRÉS DANS CE CAS

● ACF-01	● ACF-05	● 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

Waymo, a subsidiary of Alphabet (Google), operates **fully autonomous** robotaxis in several US cities. As with any autonomous vehicle, driving is delegated 100% to the agent: it perceives, decides, and acts alone, continuously. In late 2023, several Waymo vehicles were involved in **low-speed collisions** with fixed objects (barriers, poles, chains), with no injuries, but revealing a software weakness when faced with certain stationary obstacles. Waymo reported the incidents to the authorities, then filed a **voluntary software recall** in February 2024 covering around 1,200 robotaxis, with a corrective update.

Read through the lens of ACF, this case is the **reference** counterpart to Cruise: both deal with the same object, the **kill switch**, by opposite paths. Where Cruise underwent a shutdown imposed by the regulator ten days after an injury-causing accident, Waymo **detected, reported, stopped, and corrected on its own initiative**. The value of ACF here is **demonstrative**: this case shows what an internal shutdown mechanism, armed and transparent, looks like when it works.

2. Context

Waymo operates, like Cruise, a fleet of robotaxis in real conditions, at the highest level of autonomy: no human in the loop at the moment of the driving act. The concept at stake is identical to the Cruise case, but the company's behavior differs. The governance question is the same, and it remains head-on: **who stops the agent when something goes wrong, and at what moment?** This case shows that, with comparable technology, it is the governance response, not the performance of the model, that makes the difference.

3. Public sources used

- Waymo, "Voluntary recall of our previous software" (official blog, February 2024).
- TechCrunch, articles on Waymo robotaxi recalls (2024).
- InsideEVs, "Waymo Recalls 1,212 Robotaxis Because They Keep Crashing Into Things".
- The Robot Report, "Waymo recalls robotaxi software after safety failures".

4. Reconstructed AI architecture (from public sources)

Waymo autonomous robotaxi (in operation in several US cities, facts from late 2023 to early 2024).

- Nature: fully autonomous vehicle, with no safety driver on board, on public roads.
- Function: driving delegated 100%, perception, decision, and action continuously.
- Autonomy level: the highest, on a high-risk activity.

The weak signal. In late 2023, several vehicles entered **low-speed collisions** with fixed objects (barriers, poles, chains). No injuries, but a recurring pattern appeared: a software weakness when faced with certain stationary obstacles. The important point is not the isolated incident, it is the **repeated and detectable nature** of the behavior.

The company's response. Rather than waiting, Waymo **reported the incidents to the authorities**: the Phoenix police and the Arizona Department of Public Safety on December 11, 2023, then the NHTSA (federal agency) on December 15, 2023, with several follow-up exchanges through February 2024. In February 2024, Waymo filed a **voluntary software recall** covering around 1,200 robotaxis and deployed a corrective update. The company presented this recall as reflecting its responsibility to deploy its technology safely and to communicate transparently with the public.

Safeguards present (what the facts reveal):

- an internal **detection** of a recurring pattern before any serious accident;
- a **proactive report** to the authorities, without being compelled to;
- a **voluntary recall** of the fleet, decided and executed by the company;
- a **software correction** deployed to address the cause, with public communication.

5. ACF mapping

Tools used: ACF-06 (kill switch) · ACF-01 (decision map) · ACF-05 (supervision) · ACF-08 (register).

In ACF language, the response reads as follows:

ACF element	In the Waymo deployment
Named human (accountable)	Waymo, the company. It is the one that answers for the driving of its agent and that triggered its shutdown.
Agent	The autonomous robotaxi, which perceives, decides, and acts alone.
Platform	The underlying autonomous driving system (perception, planning, control).
Delegated decision	Continuous driving, under supervision of a shutdown and correction mechanism armed internally.
Autonomy level	Fully autonomous on driving, but with a reversible fleet shutdown held by the company.
Non-delegable zones	The decision to suspend and correct remains with the company: it neither delegated it nor left it to a third party.
Rule / safeguard	Voluntary recall armed and triggered internally; software correction deployed.
Control third party	The regulator (police, Arizona DPS, NHTSA), informed by the company , not forced to impose the shutdown.
Traceability	Dated report to the authorities and public communication of the recall: the incident is documented and accounted for.

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

6. Analysis

Tools used: ACF-06 (shutdown) · ACF-05 (supervision) · ACF-01 (mapping) · ACF-08 (register).

What the case demonstrates (the mechanism when it works):

- **An internal kill switch, armed and triggered.** Waymo did not wait for a serious accident or a regulatory constraint. The trigger (a recurring faulty behavior) was enough to decide, internally, a fleet recall. This is the very definition of a governance shutdown mechanism: internal, armed, decided by the company.
- **Supervision served to arm the shutdown, not only to measure.** The pattern of repeated collisions with fixed objects was detected early, before it could escalate. Useful supervision does not merely observe performance: it triggers the correction.
- **Transparency is a part of the mechanism.** The proactive report to the authorities and the public communication of the recall are part of governance, not a communications supplement. Without a report, there is no correction loop and no trust.

What ACF names and generalizes:

- a **shutdown mechanism** (ACF-06) whose trigger is a **recurring and measurable** faulty behavior, decided by the company, executed through report, recall, and update;
- a **supervision** (ACF-05) detecting a pattern early by obstacle category and **arming the recall** before the serious accident;
- a **register** (ACF-08) time-stamping the incidents, the reports, and the correction, a condition of accountability.

A voluntary recall is not an admission of weakness: it is the **proof that the governance mechanism works**. The company detects, owns, stops, and corrects. The absence of a recall is not a sign of quality, it is often a sign of blindness.

7. ACF toolkit used

This grid indicates the tools that the ACF analysis uses (✓) to read 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-06 · Kill switch	✓	Name the internal shutdown mechanism, armed and triggered by the company
ACF-05 · Supervision	✓	Detect a recurring pattern early and arm the recall
ACF-01 · Decision map	✓	Locate the agent, the accountable party, and the shutdown decision
ACF-08 · Decisions register	✓	Time-stamp incidents, reports, and correction
ACF-03 · Agentic constitution	→ _{SOON}	Enshrine the shutdown and correction rule as an obligation
ACF-02 · Criticality matrix	→ _{SOON}	Classify obstacles and behaviors by criticality

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

1. What distinguishes an **internal** kill switch, triggered by the company, from an **imposed** kill switch, forced by a regulator?
2. Does our supervision merely measure performance, or does it truly arm the shutdown and the correction?
3. From what recurring pattern do we trigger a recall, and **before** what type of accident?
4. Is transparency toward the authorities and the public part of our mechanism, or do we treat it as a supplement?
5. Do we consider a voluntary recall an admission of weakness or a proof of governance maturity?

9. General lessons

With comparable technology, two companies can produce two opposite forms of governance. Waymo and Cruise both saw their agent act wrongly. What separates them is not the performance of the model, it is the way of **stopping** the agent: Waymo detected, reported, recalled, and corrected on its own; the authority was informed, not forced to intervene. This is exactly the posture that ACF recommends.

The transferable lesson: the governance of an agent is judged by the way you stop it, not only by the technology that makes it move forward. An internal, fast, and transparent shutdown is the sign of a mature governance, not of a weakness.

10. What ACF would do (reference design)

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

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

- **define** a shutdown mechanism (ACF-06) whose trigger is a recurring and measurable faulty behavior, decided internally, executed through report, recall, and update;
- **arm** a supervision (ACF-05) capable of detecting a pattern early by incident category and of triggering the correction before the serious accident;
- **map** driving on the autonomy scale (ACF-01) while keeping with the company the decision to suspend and correct;
- **enshrine** in a signed agentic constitution (ACF-03) the obligation to report and correct, transparency included;
- **document** a decisions register (ACF-08) time-stamping incidents, reports, and corrections.

Here, ACF does not correct a failure: it **names, generalizes, and makes auditable** what a mature deployment already does, so that other organizations can reproduce it by design rather than by isolated virtue.

11. Similar cases (transferability)

This case is specific neither to Waymo nor to robotaxis. It is transferable to any organization that delegates to an agent a **continuous high-risk action** and that must be able to stop and correct it on its own initiative:

-  autonomous vehicles and mobility
-  industrial and logistics robotics
-  drones and tele-operated machines
-  automated trading and execution
-  action agents on critical systems (energy, networks)
-  medical devices and hospital automation

Everywhere, the same pattern: an agent that acts alone, a detectable failure pattern, and a shutdown and correction mechanism that the company must hold internally, armed and transparent. 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 an **official communication from the company** (documented voluntary software recall, Waymo blog) and on coverage by reference specialized press. The recall constitutes a formal act, dated and verifiable by the reader. The ACF analysis bears on **governance**, not on the internal technical details of the driving software, which are not public and have no bearing on the reading proposed here.

Teacher guide

Pedagogical objective. Convey the difference between an internal kill switch, triggered by the company, and an imposed kill switch, forced by a third party, and why the difference lies in governance, not in technology.

Session outline (90 min).

1. (15 min) Reading of the **two** cases of the kit: Cruise (incident) and Waymo (reference), without the analyses.
2. (20 min) In groups: describe each company's kill switch (card ACF-06), trigger, who decides, execution.
3. (20 min) Debate: "Is a voluntary recall a weakness or a maturity?" (expected answer: a governance maturity).
4. (20 min) Projection: which supervision indicators trigger a recall before the serious accident?
5. (15 min) Synthesis: governance is judged by the way you stop an agent, not only by the technology that makes it move forward.

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-005)

From this case, the student produces:

- the definition of Waymo's kill switch (card ACF-06): recurring trigger, who decides, execution;
- the table of supervision indicators (card ACF-05) arming a recall before the serious accident;
- the one-page comparison Cruise / Waymo: who stops, when, with what transparency;
- the schema of a decisions register (card ACF-08): incidents, reports, corrections to time-stamp.

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 decisions in Waymo's deployment and set, for each, the autonomy level. 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

Driving in normal conditions

STAKE

Perceive, brake, avoid obstacles

IMPACT

Reversible, no damage

DEADLINE

Continuous

AUTONOMY

Autonomous

B

Reporting a recurring pattern

STAKE

Alert authorities to a detected faulty behavior

IMPACT

Documented, reversible

DEADLINE

On detection

AUTONOMY

Supervised, decided by the company

C

Suspend and fix the fleet

STAKE

Decide a voluntary recall

IMPACT

Commits the company, held internally

DEADLINE

-

AUTONOMY

NON-DELEGABLE

Reading: driving is autonomous, but the decision to suspend and fix stays with the company. It is this setting that separates an **internal** stop from an imposed one.



OBJECTIVE

Link the detection of a recurring pattern to arming the recall and fix. Card filled with the public elements of the case.

1 SUPERVISION INDICATORS

Indicator	Value / target (public sources)
Collision pattern by obstacle type	detected early, before a serious accident
Report to authorities	proactive: Phoenix PD, Arizona DPS, NHTSA (Dec. 2023)
Recall scope	~1,200 robotaxis
Fix deployment	software update across the fleet
Public communication	documented recall (Waymo blog, February 2024)

2 MECHANISM

- 1 Detect a pattern early by incident category and arm the recall (ACF-06)
- 2 Supervision owner: the DDAO, who decides the recall
- 3 Report to authorities and communicate publicly, transparency included
- 4 Periodic review of obstacle and behavior categories

Here supervision does not merely measure: it **arms the fix**. Linked to the register (ACF-08), it makes the device auditable and reproducible by design, not by isolated virtue.



KILL SWITCH

Conditions for stopping and handing over to a human

LEVEL
USE
CASE
FILLED

Foundation
Control
ACF-CS-005
Waymo

OBJECTIVE

Define the signals that trigger a fleet recall and fix, decided internally. Card filled with the triggers relevant to the case.

1 TRIGGERS

Signal	Threshold	Action
Pattern of collisions with fixed objects	Recurring and detectable	Internal decision to recall the fleet
Confirmed faulty behavior	Before any serious accident	Report to authorities, corrective update
Software weakness on a stationary obstacle	Identified	Fix deployed across the fleet

2 PROCEDURE

- 1 Who decides the stop: the company (DDAO), without waiting for a regulatory constraint
- 2 How: voluntary recall, proactive reporting, deployment of an update
- 3 Regular testing of recurring-pattern detection
- 4 Log every report and recall (ACF-08)

It is precisely this device, armed on a recurring, measurable pattern, that triggered the voluntary recall of about 1,200 vehicles in February 2024, **before** any serious accident.

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



OBJECTIVE

Record incidents, reports and fixes to be able to reconstruct them, correct them at scale and account for them. Card filled with the fields relevant to the case.

1 FIELDS TO RECORD

Field to record	Example, Waymo case
Timestamp	Late 2023 to February 2024
Incident	Low-speed collision with a fixed object, no injury
Report	Phoenix PD and Arizona DPS (Dec. 11), NHTSA (Dec. 15, 2023)
Decision	Voluntary software recall (~1,200 vehicles)
Fix	Update deployed across the fleet

2 USE OF THE REGISTER

- 1 Detect a recurring pattern before a serious accident
- 2 Correct at scale via an update across the whole fleet
- 3 Establish who accounts for the agent's conduct: the company
- 4 Provide a dated trace of the report and the fix

The dated report and public communication of the recall turn the register into proof of accountability. A voluntary recall is not an admission of weakness: it is proof that the device **works**.

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

Disclaimer

This study is carried out **exclusively from public information** (official company communication and accessible press coverage). It constitutes **neither an audit, nor a certification, nor an official evaluation** of Waymo or Alphabet, nor an opinion on legal responsibilities, which fall to the competent authorities. 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

- Waymo, *Voluntary recall of our previous software* (official blog, February 2024)
- TechCrunch, articles on Waymo robotaxi recalls (2024)
- InsideEVs, *Waymo Recalls 1,212 Robotaxis Because They Keep Crashing Into Things*
- The Robot Report, *Waymo recalls robotaxi software after safety failures*

ACF Case Study ACF-CS-005 · v1 · working document, to be reviewed before any public distribution.