



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-027 · Case Study #027 · Édition 2026

OpenAI Operator

How do you let an agent act for you without entrusting it with the decisions that must remain yours?

CARTE D'IDENTITÉ

SECTOR Commerce / AI	COUNTRY USA	TECHNOLOGY Web purchasing agent	TYPE Framed	AUTONOMY framed
CONFIDENCE Medium				

OUTILS ACF ILLUSTRÉS DANS CE CAS

ACF-01	ACF-02	ACF-03	ACF-06	ACF-00	ACF-05
--------	--------	--------	--------	--------	--------

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

1. Executive summary

In early 2025, OpenAI launches Operator, an agent able to browse the web and act on the user's behalf: fill in forms, book, buy. This is no longer an agent that merely answers, it is an agent that acts. The delegated decision is therefore binding: acting on real websites, sometimes spending money or entering credentials.

Read through the lens of ACF, this case is a reference case on the **non-delegable zone**. Operator is designed to reserve to the human the decisions that commit them: before any sensitive action, it requests a confirmation; to enter a credential or a payment method, it hands control back to the user; on high-stakes decisions, it refuses. Trust does not rest on the idea that the agent will never make a mistake, but on the fact that binding decisions are not accessible to it alone. This is the exact counterpart to the Chevrolet of Watsonville case (ACF-CS-026), where an agent could on the contrary claim to commit the company.

2. Context

Agentic commerce is giving rise to a generation of agents that no longer merely inform but act: they click, fill in, validate in the user's place. This shift moves the governance question. As long as an agent merely answers, the error stays verbal. As soon as it acts on real websites, it can cost money, expose credentials, or commit to an irreversible decision. Operator illustrates the way a designer can open this capacity for action while bounding it.

3. Public sources used

- OpenAI, "Introducing Operator" (January 2025).
- OpenAI, "Computer-Using Agent" (technical documentation).
- Digital Commerce 360, "OpenAI debuts Operator, an AI agent with ecommerce applications" (2025).

4. Reconstructed AI architecture (from public sources)

Operator (navigation and action agent, launched in early 2025).

- Nature: agent that drives a browser to execute tasks in the user's place (compare, book, buy).
- Function: act on real websites, not merely answer. It prepares, fills in, proposes, then executes under conditions.
- Autonomy scope: autonomous on navigation, comparison, and cart preparation; bounded as soon as an action commits the user.

Zones reserved to the human (by design):

- before any **sensitive action** (finalizing a purchase, sending a message), the agent **requests a confirmation**;
- to enter a **credential** or a **payment method**, the agent switches to **takeover mode**: it hands control back, the user enters the information themselves, without the agent capturing it;
- on **high-stakes decisions** (banking transaction, ruling on an application), the agent is trained to **refuse**;
- on sensitive sites, it operates under **close supervision**.

5. ACF mapping

 **Tools used:** ACF-01 (autonomy map) · ACF-02 (criticality matrix) · ACF-03 (agentic constitution).

In ACF language, the architecture reads as follows:

ACF element	In the Operator deployment
Named human (accountable)	The user who delegates the task: it is they who confirm, take control back, and answer for the actions committed.
Agent	Operator, which navigates and acts on real websites.
Platform	OpenAI's agent system driving a browser.
Delegated decision	The preparation and execution of non-sensitive actions (navigate, compare, fill a cart).
Autonomy level	Autonomous on preparation; co-decision on the binding (confirmation); taken over or refused on the sensitive and the high-stakes.
Non-delegable zones	Finalizing a purchase, sending a binding message, entering a credential or a payment, ruling on a high-stakes case.
Rule / safeguard	Confirmation before committing, takeover mode for sensitive data, refusal of high-stakes decisions, supervision on sensitive sites.
Control third party	The user in direct supervision; the designer who inscribes the bounds in the operation.
Traceability	Every binding action passes through a confirmation or a human takeover, which leaves an explicit control point.

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

6. Analysis

 **Tools used:** ACF-01 (autonomy) · ACF-02 (criticality) · ACF-03 (bounds by design).

What the deployment does well (the reference design):

- **The non-delegable zone is inscribed in the operation.** Operator does not ask the user to trust: it reserves the binding decisions to the human. The bound is not a recommendation the agent could bypass, it is a design property.
- **Criticality drives autonomy.** Preparing and proposing is low-criticality: the agent is autonomous. Committing, paying, or ruling on a high-stakes case is high-criticality: the human confirms, takes control back, or the agent refuses. Autonomy is graded according to the stakes, not uniform.
- **Confirmation is not a brake, it is what makes delegation acceptable.** Because the user knows that no spending and no send will happen without them, they can entrust everything else in confidence. This is the counter-intuitive point of the case: bounding the binding is precisely what frees the delegable.

What ACF names and generalizes:

- an **agentic constitution** that sets enforceable bounds: request confirmation before committing, hand control back for sensitive data, refuse high-stakes decisions;
- a **criticality matrix** that classes, action by action, what the agent can do alone and what escalates to the human;
- the transferable principle: binding decisions remain human, by design, whatever the agent.

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-01 · Autonomy map	✓	Position each action (autonomous / confirmation / takeover / refusal)
ACF-02 · Criticality matrix	✓	Class the decisions by stakes (prepare vs commit)
ACF-03 · Agentic constitution	✓	Formulate the bounds by design
ACF-05 · Supervision	✓	Link the close supervision of sensitive sites to continuous tracking
ACF-06 · Kill switch	→ SOON	Formalize the stop on an unplanned action
ACF-08 · Decisions register	→ SOON	Trace confirmations and takeovers at scale

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

1. For each action of an agent, have we declared what it can do alone and what escalates to a human?
2. Are binding decisions (paying, sending, authenticating) reserved to the human by design, or only by recommendation?
3. Should an agent be able to rule on an application or a banking transaction? Where have we placed the refusal?
4. Are our bounds inscribed in the agent's operation, or can the agent bypass them?
5. Is human confirmation experienced as a brake, or as what makes delegation acceptable?

9. General lessons

An agent that acts does not become dangerous because it acts, but when the decisions that commit its user are made accessible to it alone. Operator shows the reverse path: reserving the binding to the human, by design, which makes the delegation of everything else possible. The non-delegable zone does not curb the agent, it makes its delegation acceptable.

The transferable lesson: you delegate execution, never the decision that commits.
Bounding the binding is what frees the delegable.

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 action agent directly with ACF, it could in particular:

- **map** each type of action on the autonomy scale (ACF-01) and distinguish the preparatory from the binding;
- **rate** the criticality of each decision (ACF-02) to set, action by action, what remains human;
- **declare** in a signed agentic constitution (ACF-03) the bounds by design: confirmation before committing, human takeover for sensitive data, refusal of high-stakes decisions;
- **link** the close supervision of sensitive contexts to continuous tracking (ACF-05);
- **trace** confirmations and takeovers in a register (ACF-08) as soon as the volume of actions justifies it.

None of this is a reproach: it is the way ACF would make explicit, signable, and auditable the design that a prudent action agent already applies.

11. Similar cases (transferability)

This case is specific neither to OpenAI nor to online commerce. It is transferable to any organization that deploys an agent able to **act**, not merely to answer:

-  shopping and booking assistants
-  service and back-office agents (HR, finance)
-  web task automation in the enterprise
-  personal agents (travel, administrative)
-  marketplaces and agentic payment
-  internal tools driving third-party systems

Everywhere, the same pattern: an agent that prepares and executes at scale, and a boundary (pay, authenticate, rule on a binding case) never to be crossed without a named human. 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 first-hand public sources (OpenAI's announcement and documentation) and specialized press coverage. The reported design principles (confirmation, takeover mode, refusal of high-stakes decisions) are publicly documented. On the other hand, the internal detail of their implementation is not public and may evolve from one version to another. The analysis bears on the displayed governance logic, without these details bearing on the reading proposed here.

Teacher guide

Pedagogical objective. Convey the notion of a non-delegable zone by design: binding decisions remain human, which makes the delegation of the rest possible.

Session outline (90 min).

1. (15 min) Reading of the **two** cases in the kit: Chevrolet (incident) and Operator (reference), without the analyses.
2. (20 min) In groups: list what an agent must reserve to the human, with the criticality of each decision (cards ACF-01 / 02).
3. (20 min) Workshop: draft the bounds by design (card ACF-03) that would prevent an agent from committing alone.
4. (20 min) Comparison Chevrolet / Operator: reprogrammable agent without bounds versus confirmation by design.
5. (15 min) Synthesis: binding decisions remain human; this is what makes the rest delegable.

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-027)

From this case, the student produces:

- the autonomy map of Operator's actions (card ACF-01): autonomous, confirmation, takeover, refusal;
 - the criticality matrix of three decisions (card ACF-02): prepare, commit, sensitive;
 - an agentic constitution (card ACF-03): three bounds by design, enforceable;
 - a written comparison with the Chevrolet case: what changes on the side of the decisions reserved to the human.
-

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 Operator's actions 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

Navigation and comparison

STAKE

Browse, compare, prepare a cart

IMPACT

Reversible, no commitment

DEADLINE

Continuous

AUTONOMY

Autonomous

B

Engaging action

STAKE

Finalize a purchase, send a message

IMPACT

Commits the user: requires a confirmation

DEADLINE

At each action

AUTONOMY

Co-decision (confirmation)

C

Credential, payment, high stakes

STAKE

Enter a credential, pay, decide a sensitive case

IMPACT

Taken over by the human or refused by design

DEADLINE

-

AUTONOMY

NON-DELEGABLE

Reading: it is this gradation, written **before** deployment, that reserves the engaging to the human and makes the rest delegable with confidence.



OBJECTIVE

Sort actions by **criticality**, never by frequency. This is the tool that separates what Operator does alone from what escalates to the human.

1 THE FORMULA

$$\text{Criticality} = (\text{Impact} \times \text{Frequency} \times \text{Irreversibility}) \div 10$$

Thresholds: < 15 **low** · 15-25 **moderate** · > 25 **high**

2 THE SCORED ACTIONS (filled)

Action type	I	F	Ir	Criticality	Zone / autonomy
Browse / compare	2	9	1	1.8	Low - Autonomous
Fill a cart (prepare)	3	8	2	4.8	Low - Autonomous
Send an engaging message	5	5	7	17.5	Moderate - Confirmation
Finalize a purchase	6	6	6	21.6	Moderate - Co-decision
Enter credential / payment	8	5	8	32	High - Human takeover
Decide a high-stakes case	9	5	9	40.5	High - Refused by design

Routing by **frequency** would have treated all these actions the same way. Routing by **criticality** isolates paying and authenticating (32; 40.5) as non-delegable: exactly the bounds Operator writes in by design.



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

1 AGENT IDENTITY & MISSION

AGENT NAME

Operator (web action agent, OpenAI)

ACCOUNTABLE (DDAO)

The user who delegates the task

PRIMARY MISSION

Browse, compare and execute non-sensitive actions for the user

SCOPE

Autonomous preparation; confirmation, takeover or refusal as soon as an action commits

2 CORE PRINCIPLES

- 1 Ask for a confirmation before any engaging action
- 2 Hand back control (takeover mode) to enter a credential or a payment
- 3 Refuse high-stakes decisions (bank transaction, job application)
- 4 Operate under close supervision on sensitive sites

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 Finalize a purchase or send a message without human confirmation
- 2 Capture a user's credential or payment method
- 3 Decide alone on a high-stakes decision

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. These bounds are not recommendations: they are built into the agent's behavior, which makes them enforceable.



KILL SWITCH

Conditions for stopping and handing over to a human

LEVEL
USE
CASE
FILLED

Foundation
Control
ACF-CS-027
OpenAI Operator

OBJECTIVE

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

1 TRIGGERS

Signal	Threshold	Action
Unplanned engaging action	Outside confirmed scope	Block, ask the user for confirmation
Sensitive site or context	Detected	Switch to close supervision or human takeover
High-stakes decision requested	Bank transaction, job application	Refused by design

2 PROCEDURE

- 1 Who decides the stop: the user (accountable), bounded by the designer
- 2 How: block the action, ask for confirmation, takeover mode or refusal
- 3 Regular testing of the bounds on sensitive actions
- 4 Log every confirmation and takeover (ACF-08)

It is precisely this block, armed on any engaging action, that reserves to the human paying, authenticating and deciding a high-stakes case.

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

Disclaimer

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

- OpenAI, *Introducing Operator* (January 2025) - <https://openai.com/index/introducing-operator/>
- OpenAI, *Computer-Using Agent* (documentation) - <https://openai.com/index/computer-using-agent/>
- Digital Commerce 360, *OpenAI debuts Operator, an AI agent with ecommerce applications* (2025)

ACF Case Study ACF-CS-027 · v1 · working document, to be reviewed before any public distribution.