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-025 · Case Study #025 · Édition 2026

Morgan Stanley (continuous supervision)

How do you keep an AI agent alive over time without letting it drift?

CARTE D'IDENTITÉ

SECTOR	COUNTRY	TECHNOLOGY	TYPE	AUTONOMY	CONFIDENCE
Finance	USA	AI supervision	Assisted AI	supervised	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

From 2023 onward, Morgan Stanley deploys an AI assistant (backed by GPT-4) for its roughly 16,000 wealth management advisors. The agent helps retrieve information from a vast internal document base, then prepare meeting summaries. Rather than launching it and forgetting it, Morgan Stanley surrounds it with a continuous supervision system: each use case is tested (evals) before deployment, output goes through human review before any binding use, and the whole is logged within a compliance framework. Result: the agent is widely adopted (on the order of 98% of the teams concerned), while remaining under control.

Read through the lens of ACF, this case is the **"reference"** counterpart to DPD: both address the same object, the **drift and supervision of an agent over time**, by opposite paths. Where DPD let a change break its safeguards without detection, Morgan Stanley designs trust as a **system** that tests, reviews, and logs continuously. The value of ACF here is **prospective**: it names what this deployment does well and provides the language to extend it, whatever the sector.

Note: Morgan Stanley also serves as the reference for the "Accountability" kit (low autonomy, non-delegable zones held). Here, the angle is different: the continuous supervision that prevents drift. The same deployment sheds light on two distinct ACF objects.

2. Context

Morgan Stanley Wealth Management runs one of the largest networks of financial advisors in the world, in a strictly regulated sector. Like any agent deployed at scale, this assistant evolves over time: new use cases, new capabilities, updates. The question the case raises is one of longevity: how does an agent stay safe not at launch, but at each evolution?

3. Public sources used

- Morgan Stanley, press releases on the deployment of the AI assistant (2023-2024).
- OpenAI, page dedicated to the Morgan Stanley deployment.
- Press coverage (CNBC) as reported.

4. Reconstructed AI architecture (from public sources)

Morgan Stanley AI assistant (in service since 2023, extended in 2024).

- Nature: internal conversational assistant, backed by GPT-4, reserved for advisors.
- Function: retrieve and synthesize information from an internal document base, then prepare meeting summaries. It **produces material** that the advisor uses; it does not commit the client alone.
- Adoption: on the order of 98% of the teams concerned.

The supervision system (what the facts reveal):

- **Evals before deployment**: each use case is tested before being put into service.
- **Human review**: the agent's output is reviewed by a human before any use binding on the client.
- **Compliance**: logging and human supervision, aligned with regulatory requirements.
- **Scope held**: the agent assists; advice and client sends remain human acts.

The governance logic. The key point is not that the agent is "perfect", but that it is **supervised over time**: evaluated before each deployment, reviewed, logged. A change or a new capability passes through the evals before going into service, which makes a drift detectable **before** it reaches the client. Trust rests on a system, not on the hope that the agent will never deviate.

5. ACF mapping

> 🧰 **Tools used:** ACF-05 (continuous supervision) · ACF-03 (agentic constitution) · ACF-01 (mapping).

In ACF language, the deployment reads as follows:

ACF element	In the Morgan Stanley deployment
Named human (accountable)	The financial advisor for each client use; above, Wealth Management leadership.
Agent	The AI assistant (document research, summary preparation). It produces, it does not commit alone.
Platform	GPT-4 / OpenAI, in an environment controlled by Morgan Stanley.
Delegated decision	No binding decision in the strong sense. The agent **prepares**, the human **reviews and decides**.
Autonomy level	**Suggestive to co-decision**: the agent proposes, the advisor validates before any client use.
Non-delegable zones	Advice and binding client sends, kept in human hands.
Rule / safeguard	Evals before deployment, human review before binding use, logging, bounded scope.
Control third party	Internal compliance function; sector regulators.
Traceability	Logging within a compliance framework; the evals are documented and replayed at each evolution.

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

6. Analysis

> 🧰 **Tools used:** ACF-05 (evals + review + log) · ACF-03 (verified constitution) · ACF-01 (mapping).

What this deployment does well (and that ACF names):

- **Supervision is a system, not an event.** Morgan Stanley does not bet on an infallible agent: it tests it, reviews it, and logs it continuously. The difference with DPD is not the quality of the model, it is the supervision system around it.
- **Drift is detectable before the client.** A deviation appears first in the evals, before going into service. Testing each use case protects the client from unverified behavior.
- **Output stays filtered.** Human review before any binding use places a named human on the path, exactly where accountability must remain.

What ACF sheds light on for what comes next:

- **The constitution must be linked to the system that tests it.** Writing rules is not enough; they are worth something only if they are **verified** after each evolution, not only at launch.
- **Good governance does not aim for perfection.** It aims for a system that **detects and corrects** drift in time. This is a transferable principle, not a privilege of Morgan Stanley.
- **The rise in autonomy will change the picture.** If an agent one day acts without review at each step, supervision will have to be designed as such: continuous evals, thresholds, decisions register.

7. ACF toolkit used

This grid indicates the tools that the ACF analysis uses (✅) and those to anticipate on the trajectory (➡ SOON). It does not mean the company uses ACF: it shows which instruments the framework would bring.

Tool	Status	Role in this case
ACF-05 · Continuous supervision	✅	Link evals, human review, and logging into a permanent system
ACF-03 · Agentic constitution	✅	Formalize the scope and guarantee its re-testing at each change
ACF-01 · Mapping	✅	Position the agent, the accountable party, and the autonomy level
ACF-08 · Decisions register	➡ SOON	As soon as an agent acts without review at each step
ACF-06 · Kill switch	➡ SOON	To anticipate if autonomy increases

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

1. Why does testing each use case before deployment protect against drift?
2. Is supervision an event (at launch) or continuous work?
3. Must an agent be "perfect", or is a system needed that detects and corrects in time?
4. Are our constitution rules linked to a system that re-tests them after each evolution?
5. Who reviews the agent's output before any use binding on the client?

9. General lessons

An agent is not safe because it "works well": it is safe because it is supervised over time. Morgan Stanley does not bet on an infallible model, it sets up a system that tests before each deployment, reviews before each binding use, and logs what is produced. It is this system, not the raw quality of the model, that makes drift detectable before it reaches the client.

The transferable lesson: trust in an agent rests on a continuous supervision system, not on the hope that it will never deviate. Supervising is permanent work, not a box ticked at launch.

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

- **map** each use on the autonomy scale (ACF-01) and set the authorized level, keeping advice and client sends outside the agent's scope;
- **declare** in a signed agentic constitution (ACF-03) the scope and the mandatory passage through human review before any binding use;
- **install** a continuous supervision system (ACF-05): evals before each deployment, human review, logging, re-testing of the rules at each change;
- **link** the constitution explicitly to the system that verifies it, so that no evolution silently erases it;
- **prepare** a decisions register (ACF-08) and a stop mechanism (ACF-06) for the day autonomy increases.

None of this is a reproach addressed to Morgan Stanley: it is the way ACF would make **explicit, signable, and auditable** the continuous supervision that a mature deployment already practices, and would equip it for the rise in autonomy.

11. Similar cases (transferability)

This case is specific neither to Morgan Stanley nor to finance. It is particularly transferable to organizations that **keep an agent alive over time under a reliability constraint**:

- ✅ banks and wealth management
- ✅ insurance
- ✅ healthcare (hospitals, laboratories)
- ✅ industry and energy (business-support agents)
- ✅ public services and administrations
- ✅ software publishers operating evolving agents

Everywhere, the same pattern: an agent that produces at scale, permanent evolution, and supervision that tests, reviews, and logs at each change. 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 first-hand public documentation: Morgan Stanley press releases, an OpenAI publication, and reference business press coverage. The reported facts are therefore solidly established and verifiable by the reader. The ACF analysis draws on these public facts; only certain internal details of the supervision system are not public, without bearing on the governance reading proposed here.

Teacher guide

Pedagogical objective. Convey that supervising an agent is continuous work, and that trust rests on a system (evals, review, log), not on the raw quality of the model.

Session outline (90 min).

1. (15 min) Reading of the **two** cases in the kit: DPD (incident) and Morgan Stanley (reference), without the analyses.
2. (20 min) In groups: describe the continuous supervision system in ACF language (card ACF-05).
3. (20 min) Workshop: write the constitution and the re-testing at each change (card ACF-03).
4. (20 min) Comparison DPD / Morgan Stanley: uncontrolled change versus evals before deployment.
5. (15 min) Synthesis: supervising an agent is continuous work, which accompanies each evolution.

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-025)

From this case, the student produces:

- the continuous supervision system (card ACF-05): evals before deployment, human review, logging;
- an agentic constitution (card ACF-03): scope, passage through review, re-testing at each change;
- the mapping of the agent (card ACF-01): agent, accountable party, autonomy level;
- a written comparison DPD / Morgan Stanley centered on supervision.

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.

AGENTIC COMMERCE FRAMEWORK®
ACF-01

DECISION MAP

Map your decisions and their autonomy level

LEVEL	Foundation
USE	Mapping
CASE	ACF-CS-025
FILLED	Morgan Stanley

OBJECTIVE Map the uses of Morgan Stanley's AI assistant and set, for each, the autonomy level. The agent prepares, the human reviews and decides. 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 — Documentary research

STAKE	Retrieve information from the curated base
IMPACT	Reversible, sourced answers
DEADLINE	Continuous
AUTONOMY	**Suggestive**

B — Meeting notes

STAKE	Prepare notes and follow-up drafts
IMPACT	Output reviewed before any use
DEADLINE	Post-meeting
AUTONOMY	**Co-decision (human review)**

C — Investment advice, engaging client communication

STAKE	Recommend, commit the firm
IMPACT	Commits accountability
DEADLINE	-
AUTONOMY	**NON-DELEGABLE**

Reading: two uses in low autonomy, the advice non-delegable. It is this map, **re-tested at each evolution**, that keeps drift detectable before the client.

ACF® · AGENTIC COMMERCE FRAMEWORK® *GOVERN AUTONOMY, PROTECT SOVEREIGNTY.* acfstandard.com

Card ACF-01 · Decision map · completed with public case data.

AGENTIC COMMERCE FRAMEWORK®
ACF-05

SUPERVISION & GOVERNANCE

Steer the agent over time

LEVEL	Foundation
USE	Supervision
CASE	ACF-CS-025
FILLED	Morgan Stanley

OBJECTIVE Link model evaluations, human review and logging into a permanent device. Card filled with the publicly known elements of the case.

1 SUPERVISION INDICATORS

Indicator	Value / device (public sources)
Evals before deployment	Each use case tested before going live
Human review	Before any engaging use toward the client
Logging	Within a compliance frame (human oversight)
Adoption	On the order of 98% of the relevant teams
Re-test at each evolution	Evals replayed before any new capability

2 DEVICE

1. Evals: test each use case before deployment and after every change
2. Systematic human review before any engaging use
3. Logging aligned with compliance requirements
4. Supervision owner: the DDAO, to be formally designated

Supervision is a **device, not an event**. Link the evals to the register (ACF-08) as soon as autonomy rises, to cover the accountability of the decision, not only the quality of the model.

ACF® · AGENTIC COMMERCE FRAMEWORK® *GOVERN AUTONOMY, PROTECT SOVEREIGNTY.* acfstandard.com

Card ACF-05 · Supervision & governance · completed with public case data.

AGENTIC COMMERCE FRAMEWORK®

ACF-06

KILL SWITCH

Conditions for stopping and handing over to a human

LEVEL	Foundation
USE	Control
CASE	ACF-CS-025
FILLED	Morgan Stanley

OBJECTIVE Define the signals that block a use case or hand over to a human. Card filled as a prospective device, to be armed as autonomy rises.

1 TRIGGERS

Signal	Threshold	Action
Eval failure after a change	Detected	Block the use case from going live
Recurring gap in human review	Above threshold	Use case suspended, review
Out-of-scope use (advice, engaging send)	Detected	Block and human escalation

2 PROCEDURE

1. Who decides the stop: the DDAO
2. How: block the affected use case, reinforced review, human escalation
3. Regular testing of the device, before each rise in autonomy
4. Log every block (ACF-08)

Here the device is **prospective**: as long as the human reviews every engaging output, the stop stays manual. It becomes necessary as soon as the agent acts without review at each act.

ACF® · AGENTIC COMMERCE FRAMEWORK® *GOVERN AUTONOMY, PROTECT SOVEREIGNTY.* acfstandard.com

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

AGENTIC COMMERCE FRAMEWORK®
ACF-08

DECISIONS REGISTER

Record who replied what, to whom, on which source

LEVEL	Foundation
USE	Traceability
CASE	ACF-CS-025
FILLED	Morgan Stanley

OBJECTIVE Record engaging outputs and their review to account for them. Card filled with the fields relevant to the case.

1 FIELDS TO RECORD

Field to record	Example, Morgan Stanley case
Timestamp	2023-2024 (deployment then extension)
Use / session	Documentary research or meeting notes
Agent output	Sourced answer / follow-up draft
Human review	Performed before any engaging client use
Status	Validated by the advisor, reviewed in compliance

2 USE OF THE REGISTER

1. Document that every engaging output has been reviewed
2. Link the evals to the trace of decisions as autonomy rises
3. Establish who accounts for the output: the advisor, then leadership
4. Provide enforceable evidence in case of regulatory control

Evals measure the **model**; the register covers the **decision**. Linking them from the rise in autonomy closes the last blind spot of supervision.

ACF® · AGENTIC COMMERCE FRAMEWORK® *GOVERN AUTONOMY, PROTECT SOVEREIGNTY.* acfstandard.com

Card ACF-08 · Decisions register · 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 Morgan Stanley. 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

- Morgan Stanley, press releases on the deployment of the AI assistant (2023-2024)
- OpenAI, page dedicated to the Morgan Stanley deployment
- CNBC, press coverage reported on the deployment of the AI assistant

ACF Case Study ACF-CS-025 · v1 · working document, to be reviewed before any public distribution.