



LEVEL	Security
DURATION	20-40 min
USE	Emergency shutdown
TRACK	Toolkit

OBJECTIVE

Be able to **immediately shut down** an agent or agentic system in the event of drift, incident, loss of control, or critical risk.

1 AGENT CONCERNED

AGENT NAME	ID / REFERENCE
<input type="text"/>	<input type="text"/>
PRIMARY MISSION	ASSOCIATED SYSTEM / SERVICE
<input type="text"/>	<input type="text"/>
OWNER (HUMAN ACCOUNTABLE)	DATE OF CREATION
<input type="text"/>	<input type="text"/>

2 AGENT RISK LEVEL

Assessment of overall risk in case of drift (check one option).

- ☐ **Critical** — major impact on finances, customers, data, or compliance.
- ☐ **High** — significant impact on operations or reputation.
- ☐ **Moderate** — limited and reversible impact.
- ☐ **Low** — minimal impact.

3 ACTIVATION CONDITIONS

The Kill switch must be activated if any of the following conditions is met:

- ☐ Behavioral drift detected
- ☐ Decision non-compliant with the rules
- ☐ Customer complaint / critical grievance
- ☐ Loss of control / critical bug
- ☐ Other (specify):
- ☐ Defined limits exceeded
- ☐ Abnormal financial impact
- ☐ Data or access breach
- ☐ Explicit human instruction



LEVEL

Security

DURATION

20-40 min

USE

Emergency shutdown

TRACK

Toolkit

4

EMERGENCY SHUTDOWN PROCEDURE

PAGE 2/4

STEP	ACTION	OWNER	TIMING / DETAIL
1. Detect	Identify the alert signal or incident		
2. Validate	Confirm that the activation condition is met		
3. Activate	Trigger the Kill switch — critical action		
4. Isolate	Isolate the agent and cut off its access		
5. Control	Verify shutdown and monitor downstream effects		

5

AUTOMATIC ACTIONS UPON ACTIVATION

- ☐ Suspend all tasks in progress
- ☐ Block integrations / APIs
- ☐ Notify the human owners
- ☐ Revoke access and tokens
- ☐ Set critical data to read-only
- ☐ Switch to degraded mode

6

NOTIFICATIONS

Whom to notify, and through which channel? (e.g., agent owner / DDAO — Delegated Decision Agent Officer, executive leadership, DPO if data is involved, business line)

ROLE / PERSON	CHANNEL (EMAIL / SMS / SLACK / OTHER)



LEVELSecurity

DURATION20-40 min

USEEmergency shutdown

TRACKToolkit

7 POST-INCIDENT REVIEW

To be performed after every Kill switch activation.

- ☐ Root-cause analysis
- ☐ Corrective actions to be implemented
- ☐ Lessons learned and documentation
- ☐ Assessment of actual impact
- ☐ Update of rules / limits

REVIEW OWNER

REVIEW DEADLINE

8 ACTIVATION HISTORY

DATE / TIME	AGENT	REASON	ACTIVATED BY	DURATION	IMPACT	REVIEW

9 KILL SWITCH ACCESS

Who can activate the Kill switch? (at least 2 people recommended)

NAME / ROLE	CONTACT

- ☐ Method: dedicated interface
- ☐ Method: system command



LEVEL	Security
DURATION	20-40 min
USE	Emergency shutdown
TRACK	Toolkit

10 ↺ REGULAR TESTING

A Kill switch that has never been tested is an uncertain Kill switch. Schedule a periodic test under real-world conditions.

FREQUENCY (MONTHLY / QUARTERLY / SEMI-ANNUAL / ANNUAL)

LAST TEST PERFORMED ON / BY

TEST OUTCOME

NEXT TEST SCHEDULED FOR

A powerful agent without a Kill switch is a liability, not an asset.

Planning the shutdown means protecting value and trust. — See **ACF-05 Supervision** and **ACF-11 Agentic Risk Assessment**.

✓ ☑ PROTOCOL VALIDATION

AGENT OWNER (DDAO) — NAME / DATE / SIGNATURE

EXECUTIVE LEADERSHIP — NAME / DATE / SIGNATURE