



CATHA PRACTICAL AI FOR CONSTRUCTION LIBRARY

When Construction AI Goes Wrong

The Construction AI Incident Response Manual

Guide 9

Practical AI for Construction. Responsible. Outcome-driven.

Catha Construction Services | 2026 Edition

How to Use This Guide

AI incidents are operational incidents. Response should be fast, evidence-based and proportional to impact.

This manual addresses failures such as incorrect external messages, duplicate actions, bad project-record changes, data exposure, runaway cost, wrong recommendations or unauthorized tool execution.

Best use

Use this guide to create a runbook before the first production failure, not after it.

Incident Categories

Category	Examples	Immediate priority
Incorrect output	Bad recommendation, fabricated fact	Stop downstream reliance; correct affected work
Unauthorized action	Message sent, record changed, transaction attempted	Contain authority; revoke access
Data/privacy	Sensitive content exposed or mishandled	Contain, preserve evidence, notify responsible owner
Runaway automation	Loops, duplicates, repeated tool calls	Stop execution, rate-limit, deduplicate
Cost spike	Unexpected token/API/tool usage	Cap spend and identify trigger
Outage/dependency	Model/API/vendor unavailable	Fail safe; route to manual process
Control failure	Approval or policy gate bypassed	Disable affected path and investigate

The 8-Step Response

1. Detect and classify.
2. Contain: disable the smallest affected capability or access path.
3. Preserve evidence: logs, prompts, tool calls, timestamps, versions.
4. Assess scope and business impact.
5. Correct or reverse what can be safely corrected.
6. Communicate to owners, affected users and required stakeholders.
7. Identify root cause and contributing control gaps.
8. Test the fix before controlled restart; update the regression set and runbook.

Severity Matrix

Severity	Definition	Response
S1	Low impact; internal; reversible	Owner fixes and records.
S2	Material workflow disruption or repeated quality failure	Pause workflow; manager + technical review.

Severity	Definition	Response
S3	External impact, sensitive data, financial or contractual risk	Contain immediately; executive/security/legal involvement as appropriate.
S4	Severe or widespread harm, irreversible action, major breach	Enterprise incident process; full shutdown of affected authority.

Restart Gate

Check	Item	Notes
☐	Root cause understood	_____
☐	Affected access reviewed	_____
☐	Fix tested against incident reproduction	_____
☐	Regression suite passed	_____
☐	Backlog/duplicates reconciled	_____
☐	Stakeholders informed	_____
☐	Monitoring heightened	_____
☐	Rollback available	_____

Construction Applications

Preserve evidence, stop further action, determine scope, protect project/customer systems, notify the appropriate owner and restore only after the cause and control change are verified.