



CATHA PRACTICAL AI FOR CONSTRUCTION LIBRARY

Is Your Construction Data Ready for AI?

A Practical Data Readiness Guide for Contractors

Guide 4

Practical AI for Construction. Responsible. Outcome-driven.

Catha Construction Services | 2026 Edition

How to Use This Guide

AI performance is constrained by the information it receives. Many “AI failures” are actually data, access, ownership or retrieval failures.

Construction data is fragmented across estimates, drawings, specifications, RFIs, submittals, schedules, daily reports, CRM, accounting/job-cost, email, shared drives and project-management platforms. AI value depends on knowing which source is authoritative and who may access it.

Best use

Assess the specific data needed for a target workflow. Do not attempt to clean the entire company before proving a use case.

The Eight Data Readiness Tests

Test	Question	Score 0-5
Availability	Does the required information exist?	
Authority	Which source is official when sources conflict?	
Quality	Is it accurate enough for the decision?	
Completeness	Are required fields/documents consistently present?	
Freshness	How quickly does information become stale?	
Access	Can the system retrieve only what it is allowed to use?	
Structure	Can information be found and interpreted reliably?	
Traceability	Can an output be tied back to its source?	

Common Data Problems That AI Exposes

- Multiple “final” versions of the same document.
- Important knowledge trapped in individual inboxes or heads.
- CRM fields that are optional, stale or inconsistently used.
- Scanned or image-only records with weak indexing.
- Shared folders with overly broad permissions.
- Policies that contradict operating practice.
- Reference data without owners or update dates.
- No way to distinguish source facts from generated summaries.

RAG, Fine-Tuning and Plain Retrieval - When to Use What

Method	Use when	Do not use as a substitute for
Search / direct retrieval	The correct source can be found deterministically.	Basic information architecture.
RAG	The model must reason over changing internal knowledge.	Permissions, source quality or evaluation.
Fine-tuning	You need repeatable behavior/style or specialized pattern learning.	Keeping changing facts current.
Structured database/API	Exact values, transactions and authoritative records matter.	Natural-language knowledge work.

Minimum Data Package for a Pilot

Check	Item	Notes
☐	Named authoritative sources	_____
☐	Sample of normal cases	_____
☐	Sample of difficult/exception cases	_____
☐	Known bad/obsolete examples	_____
☐	Access rules	_____
☐	Update frequency	_____
☐	Data owner	_____
☐	Retention requirement	_____
☐	Ground-truth test answers	_____

Catha principle

Do not give an AI system broader access than the workflow requires. Data readiness includes the ability to say “no” to unnecessary access.

Construction Applications

Do not treat every project file as equally trustworthy. Define authority, freshness, access, retention and project/customer restrictions before retrieval or automation.