



---

# Contract Administration AI

Automating Daily Work Record Reconciliation  
to Drive Operational Efficiency

**Namyun Kim** Product Manager | AI Applications

## Manual reconciliation creates significant operational delays and financial leakage

### 2-4 HOURS

#### Per Reconciliation

Contract Administrators lose 60-100 hours annually on low-value manual data processing.

### 3-4 WEEKS

#### Reporting Lag

Significant delay between work completion and record verification impacts cash flow.

### 15% ERROR

#### Average Error Rate



#### Current Manual Workflow



Contractor Submits  
DWR



Inspector Submits  
Logs



Manual Comparison

INSIGHT: CAs are trapped in "Data Entry" roles rather than "Strategic Oversight"



## Inconsistent reporting between field and office leads to protracted disputes

Metric	Contractor Record	Inspector Record
Labor Hours	8 man-hours	37 man-hours
Equipment	2 machines	6 machines
Date Recorded	June 7, 2021	June 18, 2021



### The Challenge

Identifying which record is accurate requires hours of investigation into field notes and logs, often weeks after the event.



### Business Impact

Delayed approvals impact contractor cash flow and increase the risk of legal claims and project overruns.



---

## Direct field observation identified "Reactive Administration" as the primary pain point



---

### 2 Years Field Immersion

Direct experience as a Junior Field Inspector provided deep empathy for the workflow.



---

### 50+ Reconciliation Cycles

Observed and analyzed the end-to-end process of data collection and verification.



---

### Stakeholder Interviews

Engaged with Project Managers and CAs to identify systemic bottlenecks.

---

*“By the time issues were identified, field crews had moved to different areas and corrective actions were more expensive.”*

## Inefficiency is not just a process issue; it is a measurable financial drain

### Time Waste

**2-4 HOURS**

Per DWR reconciliation. Translates to \$3,000 - \$6,000 in direct labor waste per administrator annually.

### Operational Delay

**30-DAY LAG**

Extended reporting cycles make memory-based verification impossible and stall contractor payments.

### Error Propensity

**15% VARIANCE**

Manual typing leads to formula errors, version control conflicts, and inaccurate financial reporting.

### Verification Difficulty

**HIGH FRICTION**

Disputes are harder to resolve weeks after the event, leading to "settling" rather than accurate resolution.

**Total Annual Financial Leakage: \$6,000+ Per Administrator**

## Transitioning from manual entry to AI-assisted exception handling reduces effort by 85%

### Platform-First

Reusable components designed for horizontal scalability (1 to 100 teams).

### Governance by Design

Compliance and audit trails built into the core architecture from day one.

### User Trust

Human-in-the-loop ensures AI assists while humans retain final accountability.



#### Ingest

Upload Contractor and Inspector PDFs.



#### Extract

AI parses unstructured tables in <30s.



#### Compare

Automated side-by-side variance analysis.



#### Flag

System highlights discrepancies >5%.



#### Resolve

CA reviews only flagged items (15 min).

**RESULT: 85% Reduction in Manual Effort**



## AI implementation delivers immediate ROI and shifts the team to a proactive stance

### Processing Time

Manual

**2-4 Hours**

AI-Assisted

**15-30 Mins**

### Reporting Lag

Manual

**3-4 Weeks**

AI-Assisted

**Same-Day**

### Error Rate

Manual

**15%**

AI-Assisted

**<5%**

### Annual Cost (Per CA)

Manual

**\$3K - \$6K**

AI-Assisted

**\$500**

**85%**

Time Reduction

**6 MONTHS**

Payback Period

**PROACTIVE**

Team Stance

## A modular, format-agnostic stack enables cross-domain portability

### Layer 1: Ingestion

**Docling + PyMuPDF**

#### Universal Document Processing

Format-agnostic PDF processing with configurable table detection rules for any document type.

Teams don't need custom parsers

### Layer 2: Extraction

**Ollama + Llama 3.2**

#### Adaptable AI Extraction

Schema-driven extraction using Pydantic models and configurable prompt templates per use case.

Same pipeline, different schemas

### Layer 3: Validation

**Pydantic + Custom**

#### Reusable Quality Framework

Multi-stage validation (type, range, logic) rules configurable per domain to ensure data quality.

Governance framework is portable

### Layer 4: Audit

**SQLite + Logging**

#### Universal Storage & Audit

Standard schema for any reconciliation type with every AI decision logged for regulatory compliance.

Regulatory compliance built-in



## Choosing a Local LLM over Cloud APIs was a strategic decision for data governance

### Option A

#### Cloud API (GPT-4)

- ✓ 98% Extraction Accuracy
- ✗ \$125K/Year Estimated Cost
- ✗ Data Privacy Concerns
- ✗ Requires Constant Internet

Selected

### Option B

#### Local LLM (Llama 3.2)

- ✓ Zero Marginal API Cost
- ✓ Total Data Sovereignty
- ✓ Works Offline / On-Premise
- ⚠ 85% Raw Accuracy

### The Mitigation

Implemented a multi-stage Pydantic validation framework to catch hallucinations and formatting errors, boosting final accuracy to 95%.

### The Rationale

In construction, data sovereignty and cost-predictability are higher priorities than a marginal 3% accuracy gain from cloud providers.

## A "Human-in-the-Loop" approach ensures 100% reliability in financial records

<b>Layer 1</b>	<b>Type Validation</b> Ensures all extracted quantities (hours, units, costs) are valid numerical data types before processing.
<b>Layer 2</b>	<b>Range Validation</b> Flags "impossible" or outlier values (e.g., 200% compaction or 150°F concrete) based on historical project bounds.
<b>Layer 3</b>	<b>Logic Validation</b> Cross-references related fields for internal consistency (e.g., if status is "Fail", compliance flag must be "False").
<b>Layer 4</b>	<b>Human Review</b> Final approval remains with the Contract Administrator; AI identifies discrepancies, but humans make the final decision.

**ZERO**

False negatives in identifying critical financial discrepancies

# Scaling the solution across the organization yields significant compounding returns

Initial Investment		Annual Savings (Per CA)	
Development (One-time)	\$30,000	Labor Efficiency (75 hrs)	\$3,750
Deployment & Training	\$5,000	Error Reduction	\$1,000
Total Investment	\$35,000	Total Per Admin	\$5,250

Annual Savings

\$52,500

For a 10-CA Organization

Payback Period

8 MONTHS

Rapid break-even point

5-Year NPV

\$227,000

Net Present Value at 10%

# Successful AI products are built on deep user empathy and platform thinking

1

## Start with User Pain

Two years of field observation ensured we solved the *right* problem, moving from "Data Entry" to "Exception Management."

2

## Platform Thinking

Building reusable, domain-agnostic components enables horizontal scalability across multiple organizational use cases.

3

## Governance = Trust

Users adopt AI when they trust the outputs. Multi-layer validation and human-in-the-loop are non-negotiable for adoption.

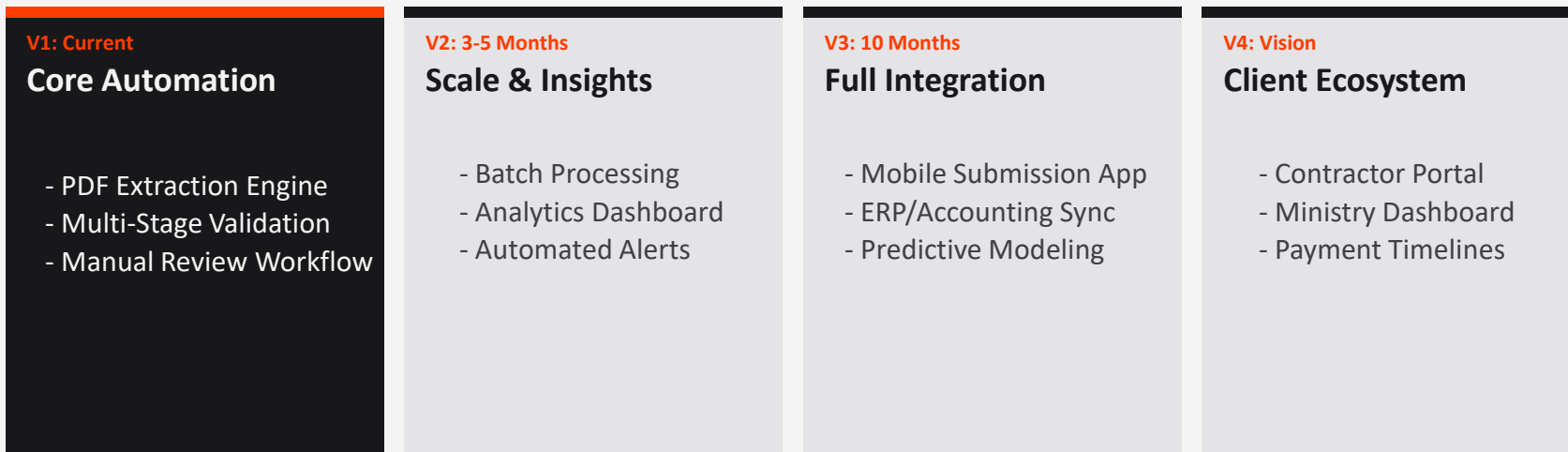
4

## Stakeholder Alignment

Success requires early buy-in across technical (IT) and operational (Compliance, Finance) teams to ensure seamless integration.

**CORE LESSON: AI is the engine, but Platform Strategy is the accelerator.**

## V1 proves the concept; V4 envisions a fully integrated field-to-office ecosystem



### Product Vision

*Eliminating manual data entry across the entire construction lifecycle to drive radical transparency and efficiency.*

## The same AI validation pattern applies to multiple high-value use cases

### Schedule Reconciliation

90% Reusable

**Comparison:** Planned vs. Actual progress

**Validation:** Date alignment, milestone completion

### Budget Tracking

85% Reusable

**Comparison:** Estimated vs. Actual costs

**Validation:** Variance thresholds, approval workflows

### Quality Control

90% Reusable

**Comparison:** Specifications vs. Test results

**Validation:** Pass/Fail logic, evidence trail

### Invoice Validation

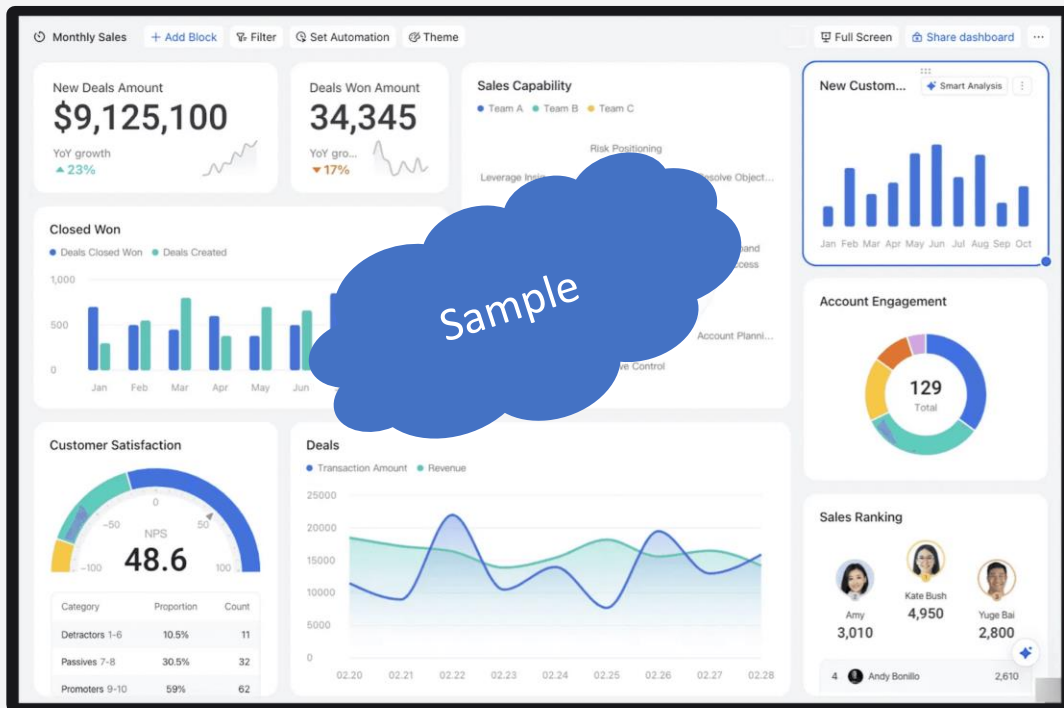
95% Reusable

**Comparison:** Payment verification vs. Approval

**Validation:** Payment verification, approval workflow

**STRATEGY:** Identifying patterns, not just solving one problem, maximizes platform ROI.

## The interface focuses user attention where it matters most: the discrepancies



### Side-by-Side View

Direct comparison of Contractor vs. Inspector data extracted from source PDFs.

### Variance Indicators

Color-coded alerts (Red > 10%, Yellow > 5%) immediately highlight critical issues.

### One-Click Drilldown

Instantly view the original document source for any flagged line item.

**Designed for Speed & Clarity**



---

***"I build **AI products** that solve  
**real-world problems.**"***

- Platform approach enables horizontal scalability
- 85% reduction in manual effort (\$55K+ savings)
- Governance-first design builds user trust



Namyun Kim



Namyun.kim@ue-Germany.de



[www.linkedin.com/in/namyun-kim-730626182](https://www.linkedin.com/in/namyun-kim-730626182)