

From Data Silos to Data Lakes

Building the Foundation for AI in HOA Management

Frank Thomas

Chief Technology Officer | Superkey Insurance · May 6, 2026 · Charlotte, NC

ABOUT THE SPEAKER



Frank Thomas

Chief Technology Officer | Superkey Insurance

Co-founded Rubikloud (acq. Kinaxis). Built data systems across retail, adtech, cloud consulting, and now HOA insurance. Waterloo Computational Mathematics. Career-long data nerd.

AREAS OF EXPERTISE

Data Architecture

AI/ML Systems

Data Strategy

Startup → Enterprise

superkey.com

frank@superkey.com

<https://ca.linkedin.com/in/frankluthertomas>

01 The Problem

Multimillion-dollar assets, managed on disconnected systems

*How many **communities** do you manage?*

How many communities do you manage?

Sales: 127

- Counts every signed contract
- Includes 4 not yet onboarded
- "They're on the board"

Operations: 119

- Only fully transitioned communities
- Staff assigned, systems live
- "We don't manage those yet"

How many communities do you manage?

Accounting: 123

- Anything they're actively billing
- Includes 2 in termination notice
- "Contracts run through June"

Website: 130

- Marketing updated 8 months ago
- Nobody told them 5 communities left
- "We should probably fix that"

Four answers, all defensible, none wrong

The biggest barrier to AI isn't the technology. It's that your data doesn't agree with itself.

Not Just Separate Systems

1. **Systems Silos:** Data in tools that are not connected
2. **Definition Silos:** Same word, different meaning
3. **Knowledge Silos:** Context that only exists in someone's head

02

Why This Matters Now

The AI Amplifier

Human Safety Net is Gone

The Old Way

- Human pulls from 3 systems
- Reconciles in Excel
- Catches contradictions
- Slow, but **reliable**

The AI Way

- AI pulls from 3 systems
- Skips reconciliation
- Misses contradictions
- Fast, confident, but **potentially wrong**



Garbage In, Garbage Out

— **Tech Nerd Proverb**

AI makes garbage sound more convincing

03

Bringing It Together

Connected is not the same as aligned

ERP

Email

Payments

Docs

PMTools

↓ *data flows in* ↓

Data Lake

Data Catalog

- **Revenue:** recognized per GAAP
- **Delinquent:** 60+ days, excluding payment plans
- **Active Community:** signed, onboarded, first billing complete

↓ *shared understanding* ↓

Dashboards

AI Assistants

Reports

People

The Missing Piece

Your ERP stays your ERP. The lake sits alongside it – adding shared definitions across all systems

Robert Chen at Riverside Gardens emailed asking about an unresolved landscaping issue. What's the full context and what should we do?

Read 4 files (ctrl+o to expand)

Here's the full picture on Robert Chen's complaint, cross-referencing multiple systems:

The Complaint

Robert Chen (Unit 14, Riverside Gardens) emailed his community manager **Sarah Chen** on **April 5** about dead hedges along the east fence line. He submitted the work order on **March 10** – meaning it's been **56 days** as of today with no resolution.

Definitions Applied

- **Work Order Response Time:** Calendar days from submission in the system to vendor arrival on site. (*Operational Definitions*)
- **Work Order Completed On Time (Routine):** Must be completed within **5 business days**. (*Operational Definitions*)

This work order is massively overdue – roughly **10x beyond the routine SLA**.

Cross-Referenced Context

Community: Riverside Gardens is an **active community** (since 2019), 180 units, \$310/month assessments, managed by Sarah Chen.

Likely Vendor Issue: Two landscaping vendors are in the system – **GreenScape Pros** (1.2-day avg response, 4.5/5 rating) and **Premier Lawn Care** (2.8-day avg response, 3.8/5 rating). Premier Lawn Care is flagged in the vendor issues log for:

- Response time increasing from 1.5 to 2.8 days since January
- Staffing issues reported
- Unexplained or possible employment

Asked about one email. AI cross-referenced the ERP, vendor performance, and board calendar

This pattern – repeated "vendor has been notified" responses with no action and an owner reporting a neighbor had the same experience – has been reported here, and the staffing problems are causing work orders to fall through the cracks.

Risk Factors

I just got back from vacation and have a ton of emails. What are the most urgent owner and board issues I need to deal with this week, prioritized by risk?

Read 3 files (ctrl+o to expand)

Here are all 8 inbound emails ranked by risk, cross-referenced against financials, insurance, vendor, and board calendar data. Today is **May 5, 2026** – several of these are already past deadline.

1. CRITICAL – Bayshore Towers: Elevator Safety & Liability

Email #8 – Linda Marks (Board VP), April 18

- Elevator #2 trapped elderly residents, fire department responded – this is a **bodily injury liability exposure**.
- 5th incident this year. Metro Elevator Services flagged in vendor report: 4 service calls in Q1 alone, 3.2-day avg response, 3.6/5 satisfaction.
- Board approved a **\$15,000 modernization study on March 20** with a 60-day timeline – that report is **due ~May 19**.
- Insurance renewal is June 30 with Chubb (\$112K premium); the policy note says **elevator condition may affect terms**.
- Linda explicitly asked that the insurer be notified – unclear if this has been done.

Action this week: Mike Torres must (1) confirm the insurer was notified of the April incident, (2) get a status update on the engineering report, (3) schedule the emergency board meeting Linda requested.

2. CRITICAL – Sunrise Palms: Recall Petition – Legal Deadline Likely Blown

Email #4 – James Walsh, April 10

- Recall petition filed **March 15**. CC&Rs require signature verification within **15 business days** (~April 3-5).
- As James' email on April 10, no acknowledgment has been sent. Today is **May 5** – the deadline is **~30 days past**.
- Elevator #10 also has elevated delinquency: **16 accounts 60+ days past due** (applying the Delinquent Account definition: 60+ days, excluding 30-day grace period). Sunrise Palms has an overall delinquency rate of **11.3% (16/142)**.

Five data sources. Prioritized by risk. No human could do this in under an hour.

Action this week: Lisa Park must immediately (1) verify petition receipt and signature count, (2) send formal acknowledgment to James Walsh with timeline, (3) consult legal counsel on the missed deadline exposure.



data



board-
calendar.md



communities.md



correspondence.
md



definitions.md



financials.md



insurance.md



vendors.md

Files in a folder. A definitions document. That's your data lake v1.

04

Monday Morning

Data is a product, not a byproduct

Treat Data Like a Product

A Product Has:

- An owners. A name, not a dept
- Users who depend on it
- A quality standard
- Documentation

Your Data Needs:

- A person responsible for each definition
- A board, investors, AI tools, managers
- Definitions reviewed, verified
- Plain English: what it means, excludes

Field	Example
Definition	Assessments collected, net of refunds
Includes/Excludes	Includes special assessments. Excludes late fees, onboarding communities
Source	General ledger — posted transactions
Owner	Jane Smith, Controller
Last Verified	March 2026

Five Fields Per Definition

Using Monthly Revenue as an example

[Source / credit if needed]

The Workshop

1. Get department leads into a room
2. Ask: What 15 terms would change a decision if defined differently?
3. Define each one, assign an owner
4. Publish somewhere, shared

You Can Start This Week

1. **Now** - Shared definitions document: shared, owned
2. **This month** - Hand definitions to AI as context. This is 80% of the way there
3. **Over time** – Formalize, connect systems, build the lake, semantic layer

Technology changes, definitions persist

KEY TAKEAWAY

***Block two hours this week. Define 15 terms.
Assign owners.***

THINK TANK HOA

THANK YOU

Block two hours this week. Define 15 terms. Assign owners.

WEB

superkey.com

EMAIL

frank@superkey.com

LINKEDIN

linkedin.com/in/franklutherthomas

DISCUSSION PROMPT

5 minutes

Turn to someone near you or
use the chat

What's your version of 'how many communities'?

Things to consider:

- What question takes your team 3+ hours to answer?
- What terms would your departments define differently?
- Where does critical knowledge live in someone's head?