

#FABCONSQLCON2026

FABCON

Microsoft Fabric
COMMUNITY CONFERENCE

SQLCON

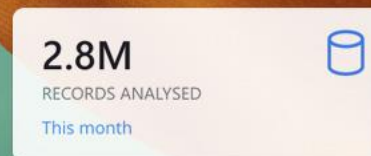
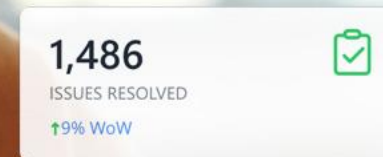
Microsoft SQL
COMMUNITY CONFERENCE

ATLANTA MARCH 16 - 20, 2026



CluedIn

Running Autonomous AI Agents in Production:
A **Komatsu** Case Study in
Fabric + Purview + AI + CluedIn



Trusted by global enterprises. Proven across industries.

From media to finance, manufacturing to risk management - leading organizations choose CluedIn to power data trust, automation, and AI-driven transformation.

MEDIA AND ENTERTAINMENT

We chose CluedIn for its powerful AI capabilities and seamless Azure-native integrations to empower our business and technical teams with data that's ready for insights & AI-driven innovation.

FELIX BAKER
HEAD OF DATA SERVICES





WEALTH MANAGEMENT

We chose CluedIn because we share a common vision of accelerating value creation through the responsible democratization of data.

PERNILLE OLLING
SENIOR VICE PRESIDENT





RISK MANAGEMENT & INSURANCE

As a Graph-based system, we have experienced firsthand how easy it is to ingest and map data, and how simple it is to define our data model.

ROBERT BROWN
ENTERPRISE ARCHITECT





SEGA®

BUCHERER
1888



amsive

Barentz.



ELEMENTIS

elitmind

Telefónica Tech

data on
Power your data

Gallagher

EVERCORE



PLAINS

ICG

We are all chasing a similar goal

How can I run AI, at scale, against my company data – without all the “magic”?

Enterprises can't use their data - they don't trust it.



88%

of companies report that their data is of little to no trust, preventing them from using it to make decisions.

Data fuels automation, AI, and insights - but broken data pipelines and fragmented systems stall progress. **Even with the best tech stacks, companies can't scale because they don't trust their data.**

**FOUNDED:**

1921

HEADQUARTERS:

Tokyo, Japan

INDUSTRY:

Construction, mining, forestry,
and industrial equipment
manufacturing

GLOBAL FOOTPRINT:

Operations in 150+ countries;
major manufacturing presence
in Asia, North America, South
America, and Europe

NEW DATA STACK:

Microsoft Fabric, Azure
Databricks, Power Platform,
Microsoft Purview, Azure
Foundry AI, CluedIn, Power BI.

The challenge, and the goal

KOMATSU[®]



CHALLENGE:

Multiple ERP systems, scale issues, inefficient processes caused by technology limitations. Limited hands-on for the project.



GOAL:

Using a Microsoft stack, provide a modern data platform with full governance to be able to run AI at scale where AI Agents do most of the data management tasks and not require large internal teams to maintain. Involve the business in being responsible for their data.



REQUIREMENTS:

- AI must be 100% controlled.
 - AI needs to work at scale, hook directly into governance for automated audit trails, lineage and explainability.
 - 10 million records a day to process.
 - Must be implementable by the business, not IT.
-



Hi, how can I help you?



Agentic Data Management with AI Agents and Copilot

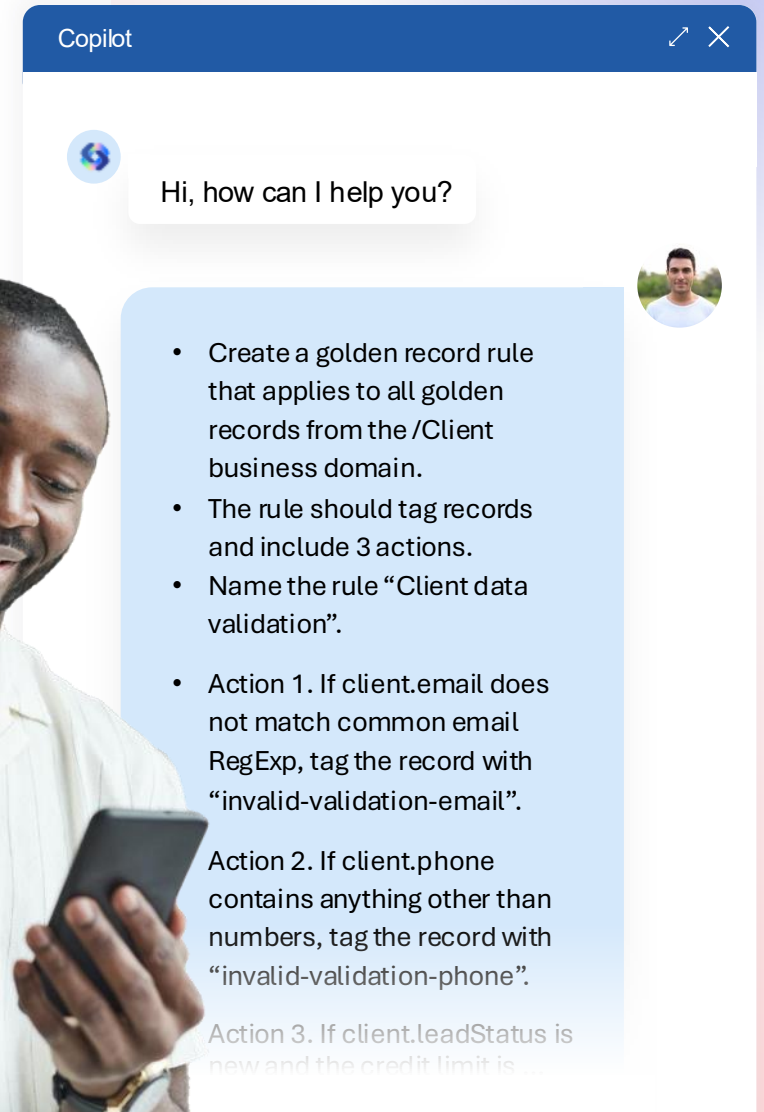
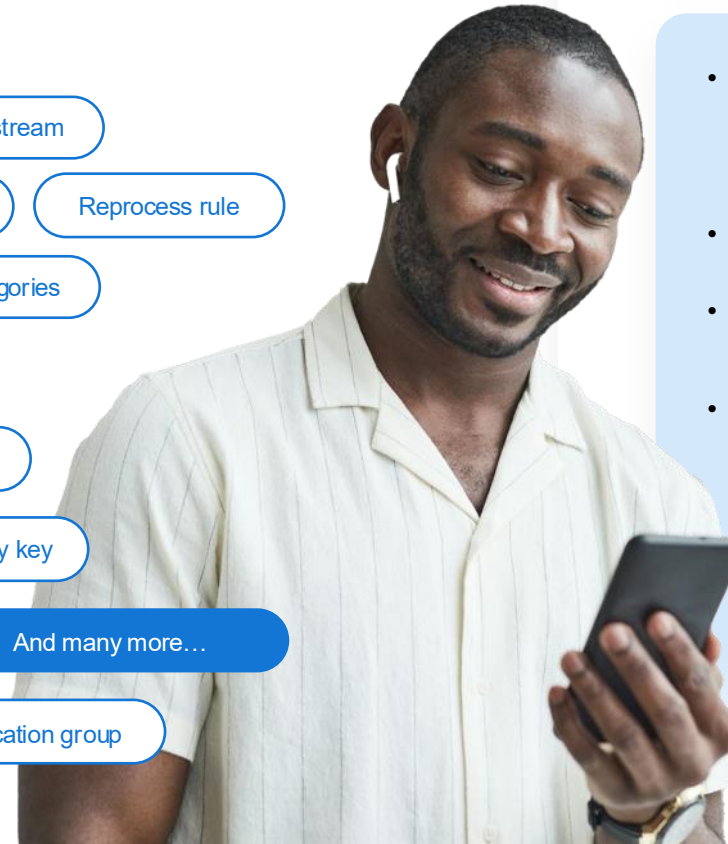
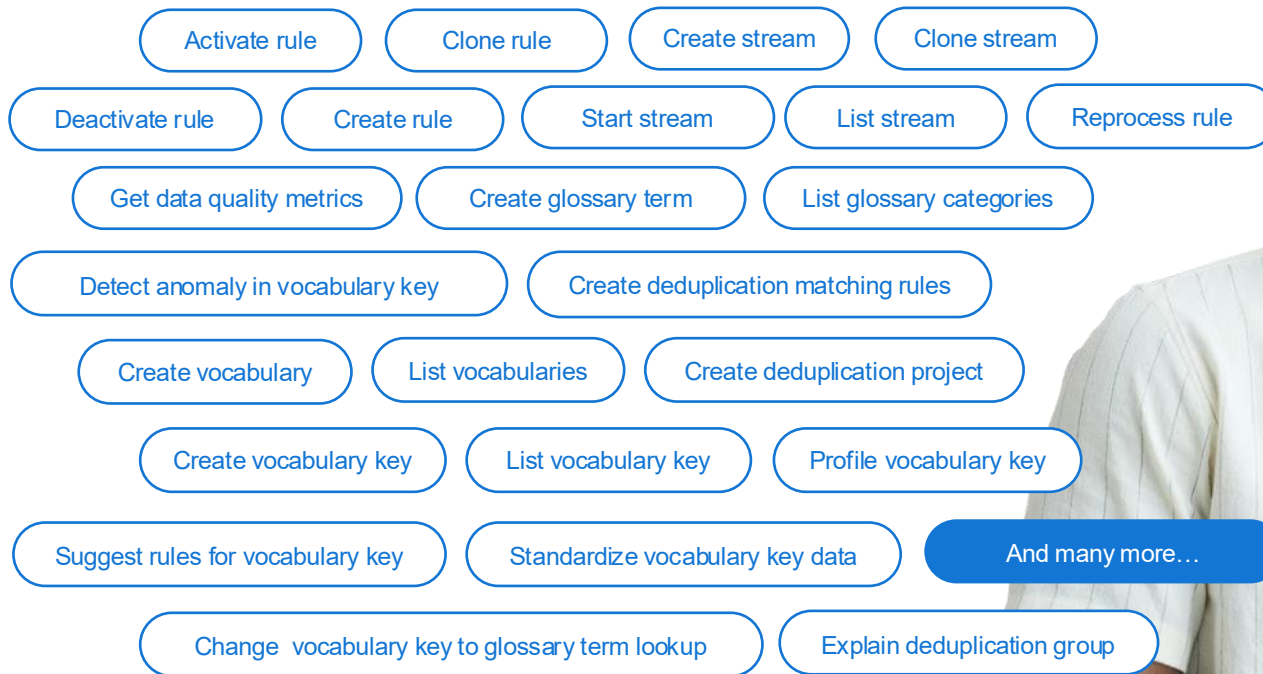


Scale the challenge of using AI....to get your data ready for AI

FROM PROMPT TO PRODUCTION

Describe it. Teach it. Let AI Agents do the work

100% Sandboxed. 100% Governed.



INTRODUCING AI AGENTS

Autonomy with control

CluedIn AI Agents have all the guardrails we know that are needed in order to allow this to actually happen in production on real, production data.

Workflow

Decisions are federated through workflow processes

Always-on intelligence

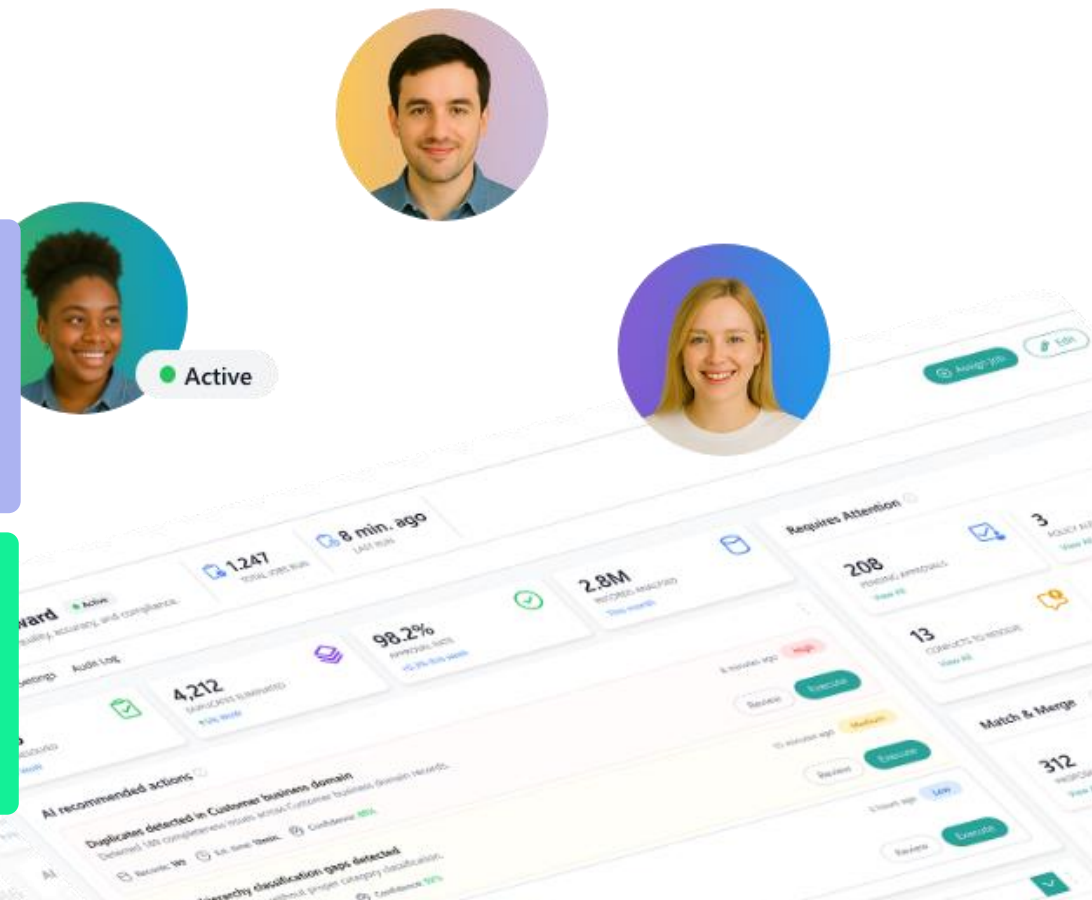
Continuous cleanup and enrichment, 24/7

Smarter decisions, faster

Agents surface insights before you ask, explained

Built with trust in mind

You stay in charge, with full transparency



The reality of running AI Agents on data at scale

It took one year of collaboration with our customers to define the guardrails required before running AI Agents safely in production.

Guardrails for Production AI

- Must be read-only by default.
- Cost-efficient at scale.
- Integrate seamlessly with enterprise workflows.
- Automatically generate deterministic audit trails, lineage, rules, and policies.

Platform philosophy

- Provide an open canvas - flexible enough for end users to solve real business problems.
- AI is significantly cheaper than many assume.

Practical reality

- Context window limitations remain a technical constraint.



Our rules on AI Agents

AI Agents must...

Governance & Control



Be read-only on master data



Adhere to access control



Provide attribution



Soft-create only



Generate deterministic audit trails and lineage



Produce results that are easy to identify and roll back

Workflow & Trust



Use the same tools as humans



Support human-in-the-loop workflows



Explain reasoning



Track deterministic data quality metrics

Autonomy & Scale



Work without prompting



Be able to iterate



Operate on millions of records

What Komatsu needed from AI agents?

KOMATSU®

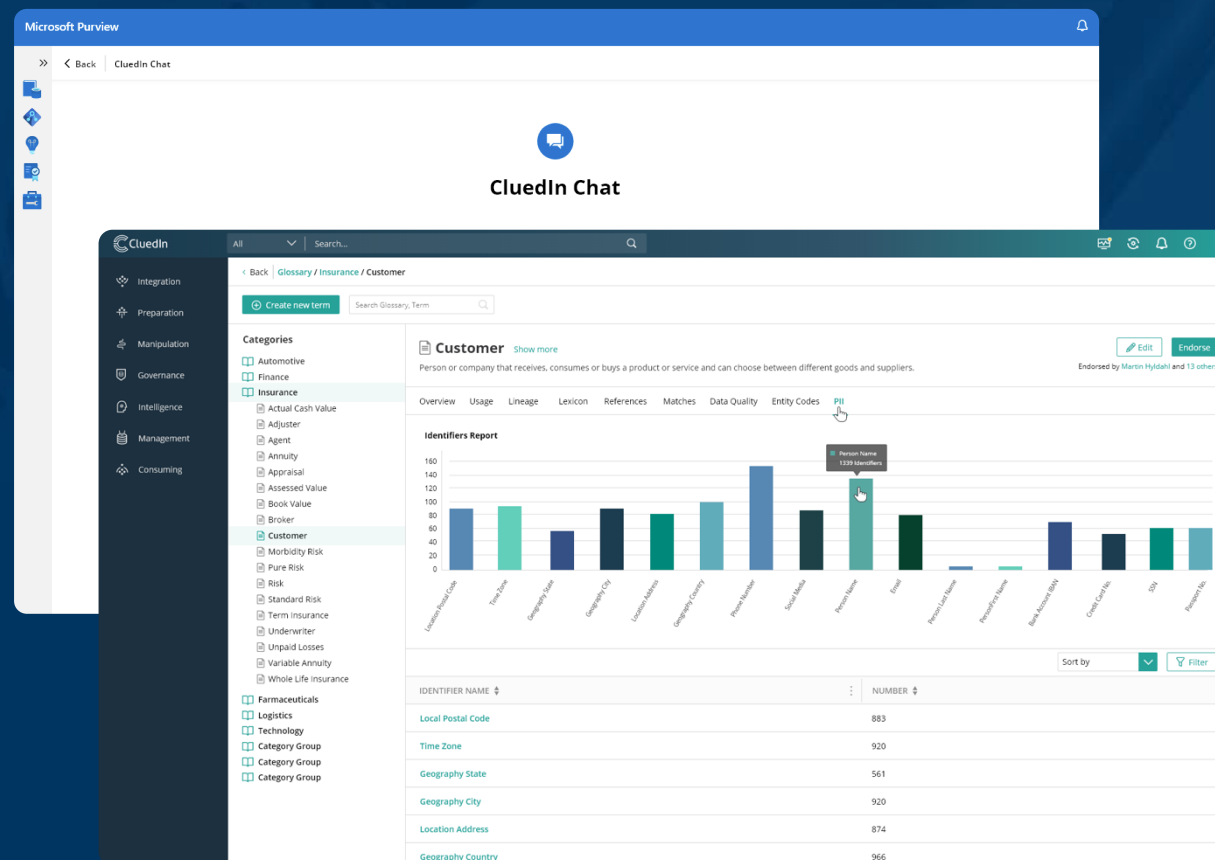
Awareness of Komatsu-specific terminology

- Abbreviations
- Data dictionary

Ability to run in Microsoft Fabric

Federation into existing working channels

- Microsoft Teams
- JIRA





Purview's Role



Near Realtime Snapshot of Dataverse, Fabric at any point in time.



Central lineage engine.



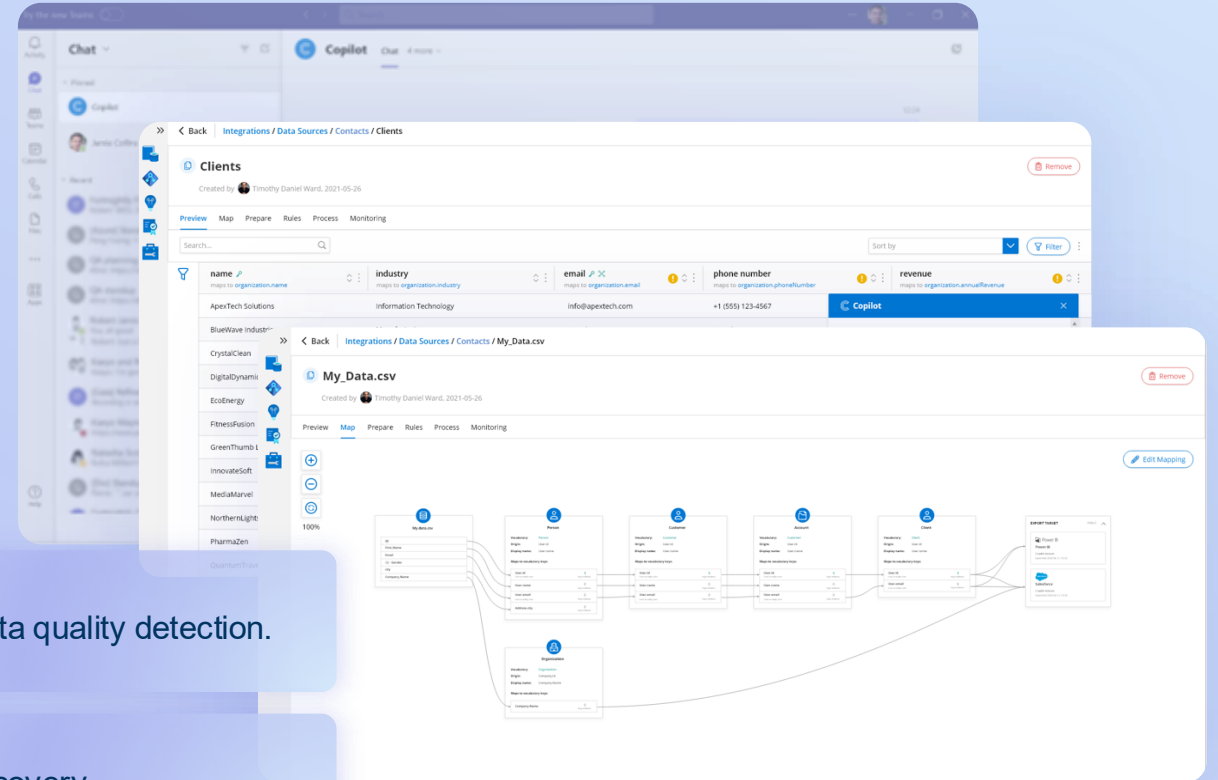
Holder of Key Vaults for access to data across the estate.



Early data quality detection.



Data discovery.





Fabric's Role



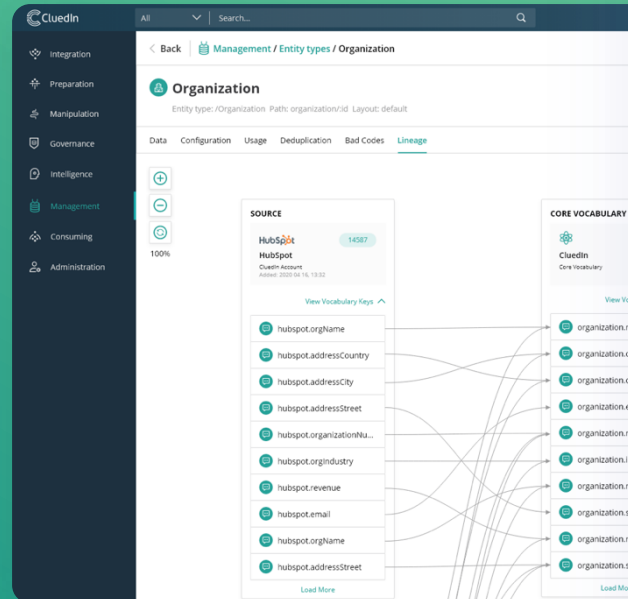
Data movement / Data integration.



Delta tables.



Analytics / Reporting.



Import data from ingestion point

1 Choose Mapping type

2 Configure

Choose mapping type

☒ Auto Mapping

CluedIn will automatically map the data set for you. You can modify this mapping once created.

☐ Manual Mapping

CluedIn will create an empty mapping configuration. You will have full control over the mapping and how each field for your data set will generate records.

☐ Existing Mapping

Reuse an already existing mapping. This is useful for when mapping the same dataset or a dataset with the same data structure.

CluedIn AI Recommendation

BETA

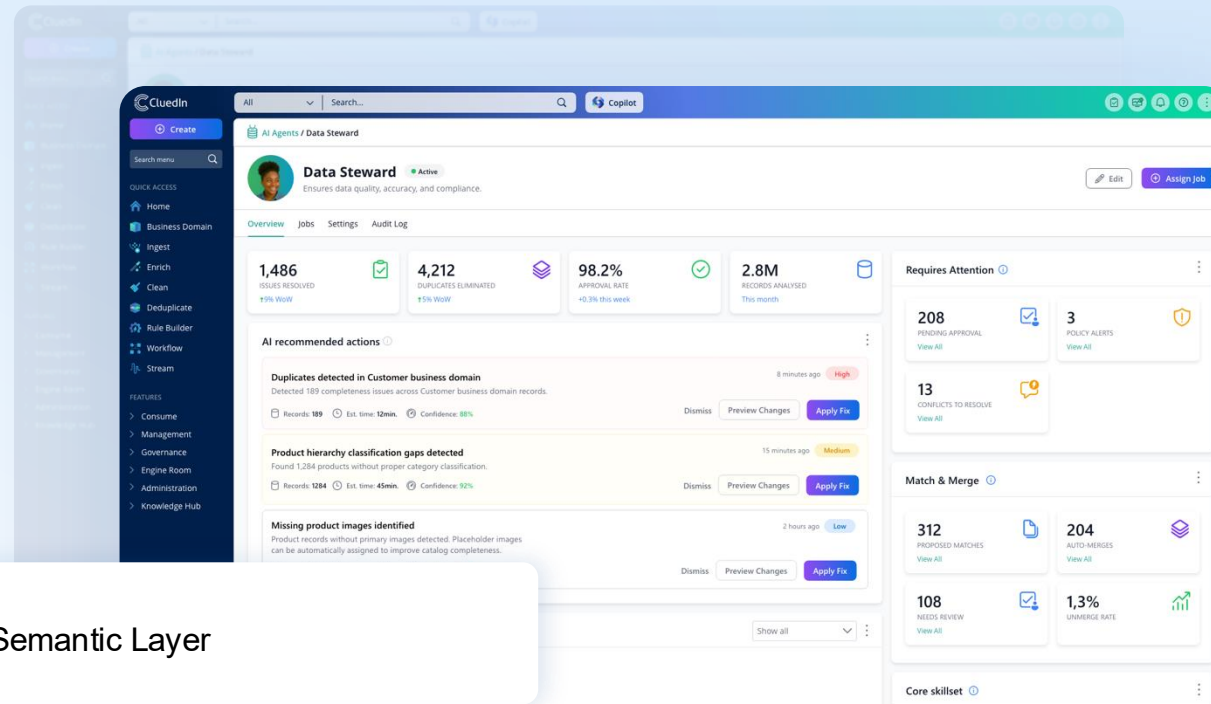
Using the AI recommendation engine can be done safely. Each decision taken by the AI engine can be modified later.

Cancel

Next



CluedIn's Role



Entity-Level Data Quality



Business Semantic Layer



Matching and Merging



AI Agent first responders for data validation, data quality, mapping, matching / merging.



Data Enrichment



Data Consolidation

And now adding even more speed to the Augmented CluedIn Platform...

The World's First Data Quality and AI Integration



Natural language processing

Give any business user the tools to instantly enrich, correct, validate and standardize data with conversational prompts in any language. No engineers needed - just tell CluedIn what you want to do, and it's done.



Risk free automation

Automate easily, but never mysteriously. Ask CluedIn AI for detailed justifications on any action, and instantly roll back any changes that you're not confident in.



Your master data co-pilot

Gain confidence in every decision you make with the industry's most powerful AI. Validate your actions and benefit from intuitive corroboration of your findings.



Immediate Data Governance

Input data policies directly into the platform and enforce them instantaneously.

What's happening in the market - and how CluedIn differs

MARKET REALITY

“Get your data ready for AI.”

Same promise made for BI/ML/RPA - hasn't delivered.

Data stewardship fatigue.

Critical but undesirable work, users avoid manual governance.

Regulation rising.

EU AI Act & global governance demand auditability and control.

Execution gap persists.

DQ projects are expensive with low success, observability only reports what's wrong.

CLUEDIN APPROACH

Use AI to get data ready.

Agentic automation prepares and maintains data for AI/BI/ML - continuously.

Agentic Data Management.

Autonomous agents do the heavy lifting with traceability, human-in-the-loop when needed.

Built-in guardrails.

Automated lineage, rollback, explanations, workflows, and audit trails - production-ready compliance.

We fix it, not just flag it.

CluedIn moves beyond observability – agents auto-remediate issues and maintain a clean data foundation.

Let's see it in action!

Using AI Agents to automatically respond to new data and manage every step of the data quality flow



I keep your data clean, consistent, and reliable.

Data Steward

Ensures data quality, accuracy, and compliance across all records, helping maintain trustworthy and reliable information throughout your systems.



I build the foundation that keeps your data flowing.

Data Architect

Designs and implements the technical blueprint for an organization's data infrastructure, ensuring data is accessible, reliable, and secure.



I find what's wrong in your data so we can fix it fast.

Data Auditor

Evaluates your data and flags records that do not comply with ISO and IEC standards, clearly explaining why, so you can easily assess and improve the overall health of your data.



I fill in the blanks to make your data more complete.

Data Enricher

Analyses your data and uses general knowledge along with online sources to fill in missing information or enrich existing records, improving the overall quality and value of your data.



I protect your data from compliance risks.

Data Compliance Expert

Tags every data record that contains information breaching global regulatory requirements such as PII, GDPR, CCPA, PCI DSS, and upcoming data protection regulation ensuring compliance and reducing risk exposure.



I turn insights into action and keep progress moving.

Team Lead

Analyses data and automatically generates new tasks based on the insights it uncovers, enabling continuous improvement and smarter decision-making.



I set the rules that keep your data trusted and compliant.

Data Governance

Ensures data quality, accuracy, and compliance across all records, helping maintain trustworthy and reliable information throughout your systems.



I'm your blank-slate Agent, built entirely around your vision.

Your AI Agent

Create custom AI Agent to fit your organisations needs.

Agent Skills:

- Map data
- Enforce validation and tag invalid records
- Match and merge duplicates at scale
- Fix data quality issues
- Enrich data by looking up data brokers / web searches
- Classify records e.g. Sensitive, Classification



Federate tasks onto channels with Power Automate



Suggested data quality fixes



Auto expansion of abbreviations



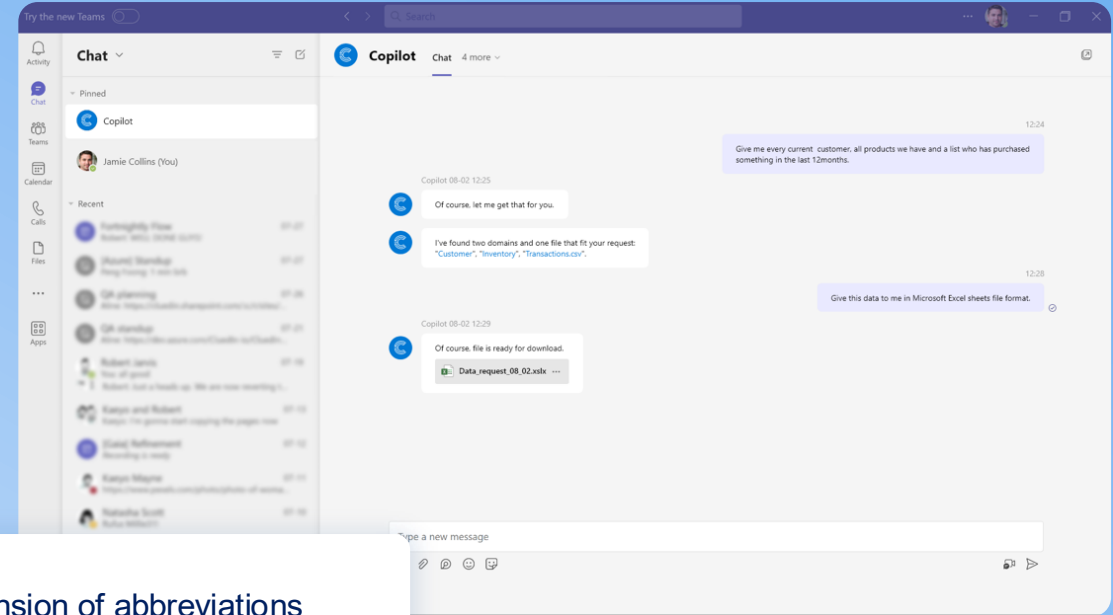
Suggested possible duplicate records



Impact analysis



Suggested new or updated validation rules



NEW CLUEDIN AI AGENTS

Create your own Data Expert Team!

Increase efficiency by a minimum of 30x
 Unmatched Support for the Business – Freeing UP IT Resources.





Always on the job



Maintaining data quality



Proactively identifying issues



Proactively fixing issues



Save time



Save money


AI Agents

Create agent

Search...

☐	Name	Description	
☐	 Sarah Data Steward	Responsible for ensuring data quality, accuracy, and proper usage by managing metadata and maintaining consistent data governance standards.	
☐	 Mark Data Matcher	Specializes in identifying, linking, and reconciling data across disparate systems to ensure accurate and unified datasets.	
☐	 Lars Data Policy Manager	Develops, enforces, and oversees data governance policies to align data practices with organizational and regulatory requirements.	 admin@foobar.com
☐	 Jane Data Product Builder	Designs and creates data-driven solutions, such as dashboards, analytics tools, or APIs, to enable effective data usage and insights.	 admin@foobar.com
☐	 Phong Data Privacy Manager	Ensures data privacy compliance by implementing and monitoring privacy frameworks, handling sensitive data securely, and mitigating risks.	 admin@foobar.com
☐	 Tess Data Domain Manager	Oversees specific data domains (e.g., customer or financial data) to ensure their integrity, accessibility, and relevance within the organization.	 admin@foobar.com
☐	 Ivan Data Auditor	Conducts thorough reviews of data practices, systems, and processes to identify discrepancies, ensure compliance, and improve data accuracy.	 admin@foobar.com
☐	 Finn Semantic Data Mapper	Focuses on aligning and translating data across systems by creating mappings that ensure consistent understanding of data meanings and relationships.	 admin@foobar.com

CluedIn

All

Search...

Copilot

Create

Search menu

QUICK ACCESS

Home

Business Domain

Ingest

Enrich

Clean

Deduplicate

Rule Builder

Workflow

Stream

FEATURES

> Consume

> Management

> Governance

> Engine Room

> Administration

> Knowledge Hub

Welcome, Alex!

Let's begin your journey towards better data insights in 2 quick steps.

Data quality by business domain

View All

8	84%	63k	1,245	954
Business Domains	Avg. Quality	Total Records	Total Issues	Auto fixable
Customer	92% ↑ +2%	12 565 Records	249 Total issues	189 Auto-fixable
Product	86% ↑ +1%	12 565 Records	249 Total issues	189 Auto-fixable
Supplier	85% → 0%	12 565 Records	249 Total issues	189 Auto-fixable
Finance	79% ↓ -2%	12 565 Records	249 Total issues	189 Auto-fixable
Employee	77% ↓ -3%	12 565 Records	249 Total issues	189 Auto-fixable
Contract	70% ↓ -2%	12 565 Records	249 Total issues	189 Auto-fixable

Recommended actions

View All

Merge 312 duplicate Customer records

AI identified patterns to fill missing email addresses using name + domain matching.

Records: 312 Est. time: 12min. Confidence: 88% Customer

Data Quality Guardian 5 minutes ago

Dismiss

Preview Changes

Apply

Recommended actions for business domain

Customer 92%

↑ +2%

12 565
Records

249
Total issues

189
Auto-fixable

Showing 15 out of 15

All action types

All impact types

Merge 312 duplicate Customer records

High Impact

AI identified patterns to fill missing email addresses using name + domain matching.

Records: 312 Est. time: 12min. Confidence: 88%

Data Quality Guardian 5 minutes ago

Dismiss

Preview Changes

Apply Fix

Validate inactive customers with recent transactions

High Impact

Some customers marked as "Inactive" still have purchases in the last 30 days. Manual review is needed to confirm whether they should remain inactive or be reactivated.

Records: 7 Est. time: 12min. Confidence: 88%

Data Quality Guardian 5 minutes ago

Dismiss

Preview Changes

Apply Fix

Standardize 1458 phone number formats

Medium Impact

Phone numbers vary in format (local, national, international). AI proposes converting all to E.164 for consistency.

Records: 1458 Est. time: 8min. Confidence: 99%

Data Quality Guardian 15 minutes ago

Dismiss

Preview Changes

Apply Fix

Correct 145 invalid email domains

Medium Impact

AI identified 145 customer emails with misspelled domains (e.g., gamil.com, hotmial.com) and suggests corrections.

Records: 145 Est. time: 8min. Confidence: 98%

Data Quality Guardian 15 minutes ago

Dismiss

Preview Changes

Apply Fix

FABCON SQLCON

ATLANTA26

JOIN THE CONVERSATION #FABCONSQLCON26

#FABCONSQLCON26

CluedIn

Create

Search menu

QUICK ACCESS

Home

Business Domain

Ingest

Enrich

Clean

Deduplicate

Rule Builder

Workflow

Stream

FEATURES

Consume

Management

Governance

Engine Room

Administration

Knowledge Hub

All

Search...

Copilot

... Data Steward / Fix data quality issues

Fix data quality issues

Created by Timothy Daniel Ward, 2 hours ago

Enabled

Total costs: \$0.03

Run Now

Test Now

Remove

Cancel

Save

Configuration

Settings

Results

Runs

Affected Records

Owners

Audit Log

1 Data

Hospitals

+ Add Data Source

2 Skills

Create data validation rule

Find duplicates

Fix data quality issues

Fix data quality issues with rules

Tag records

Instructions

Fix data quality issues in Hospitals

Preview

EntityId	Name	hospital.address	hospital.bedCount	hospital.city
00685034-d58b-50d8-8a74-d6c0ce4f4e27	Sacred Mary Hospital	228 Queen St	350	Geelong
06d66f6a-c6cb-5375-a7c7-6ccf92f8dd582	Southern George Health Campus	113 Queen St	206	ACT
07324708-ff98-43b9-9ac9-1bb2020d3f9a	St. Vincent Medical Centre	12 High St	420	Melbourne
0a94cd42-72a9-4581-bf32-26dd991dbb1	North Bay Clinic	19 Harbour Rd	75	Sydney
19fdf3c5-b0e4-4a29-9ad3-02f12a7cfa6d2	Central City Hospital	560 Elizabeth St	680	Brisbane
2b348d1f-765f-4338-a982-bc52ac45e9a4	Green Hills Community Health	78 Oak Dr	110	Hobart
32b94c2f-5a70-4765-9c0a-52c408ad3d7f	Eastside Children's Hospital	97 Garden Rd	220	Adelaide

FABCON SQLCON

ATLANTA26

JOIN THE CONVERSATION #FABCONSQLCON26

Create

Home

Ingestion

Preparation

Management

Governance

Sensitive Data

Quality Metrics

Global Data Metrics

Tag Monitoring

Workflow

Governance

Consume

Engine Room

Administration

Knowledge Hub

Tag Monitoring

Select tag to view data

Search tag...

Tag

Score

Records

Invalid Account Type

Rule: Validate Account Type Type: Data part

31.55%

2953

Invalid Billing Addresses

Rule: Check Billing Address Format Type: Data part

0.91%

85

Invalid Cities

Rule: Validate City Names Type: Data part

5.59%

523

Invalid Countries

Rule: Validate Country Codes Type: Data part

0.42%

39

Invalid Delivery Terms

Rule: Validate Delivery Terms Type: Data part

31.29%

2929

Ready for review

Rule: Golden Record Review Type: Golden record

98.8%

3289

Invalid Method Of Payments

Rule: Validate Payment Methods Type: Data part

54.94%

5142

Invalid Mode Of Deliveries

Rule: Validate Delivery Methods Type: Data part

31.22%

2922

Invalid Names

Rule: Validate Customer Names Type: Data part

1.78%

167

Ready for deduplication

Rule: Readiness for Deduplication Type: Golden record

99.9%

1726

Finance approved

Rule: Golden Record Finance Approval Type: Golden record

100%

2815

Invalid Segments

Rule: Validate Customer Segments Type: Data part

99.76%

9338

Invalid Shipping Addresses

Rule: Check Shipping Address Format Type: Data part

0.43%

40

Invalid States

Rule: Validate State Codes Type: Data part

49.35%

4619

Business Domain

Organisation

Data quality in business domains

Average % of Records Tagged

03.01 89% 03.02 79% 03.03 79% 03.04 70% 03.05 74% 03.06 69% 03.07 69% 03.08 60% 03.09 74% 03.10 67% 03.11 69% 03.12 61% 03.13 61% 03.14 61% 03.15 57% 03.16 57% 03.17 76% 03.18 55% 03.19 50% 03.20 50% 03.21 55% 03.22 69% 03.23 69% 03.24 58% 03.25 58% 03.26 45% 03.27 45% 03.28 48% 03.29 48% 03.30 43% 03.31 39%

Search record...

Column Options

Export Data

All

Name	Business Domain	Account type	Billing address	City	Country
organization.name		organization.account...	organization.billing...	organization.addr...	organization.addr...
Alpha Tech Solutions	Organisation	Corporate	123 Main St	New York	USA
Beta Logistics	Organisation	Business	456 Oak Ave	Los Angeles	USA
Gamma Retailers	Organisation	Business	789 Pine Rd	Toronto	Canada
Delta Manufacturing	Organisation	Enterprise	321 Maple Blvd	London	UK
Epsilon Finance	Organisation	Enterprise	654 Cedar Ln	Sydney	Australia
Zeta Healthcare	Organisation	Corporate	987 Birch St	Berlin	Germany
Theta Education	Organisation	Non-Profit	741 Spruce Dr	Paris	France

QUICK ACCESS

Home

Business Domain

Ingest

Enrich

Clean

Deduplicate

Rule Builder

Workflow

Stream

FEATURES

Consume

Management

Governance

Engine Room

Administration

Knowledge Hub

Search fields... 🔍

SOURCE FILES:

Customer_data.xlsx
8,421 records · 4 fields

name
e.g., John Doe

email_address
e.g., John@example.com

contact_phone
e.g., +1234567890

birth_date
e.g., 1985-03-12

id
e.g., 12345

Cust_data_25.csv
6,236 records · 4 fields

orders_2025.csv
12,874 records · 6 fields

products.json
2,156 records · 5 fields

TARGET SCHEMA:

Customer
from 2 source files

Name
customer.name
String

Email
customer.emailAddress
String

Phone Number
customer.phoneNumber
String

Birthday
customer.birthDate
Date

Id
customer.id
Integer Identifier

Registration Date
customer.registrationDate
Date New Draft

Total Orders
customer.totalOrders
Integer New Draft

Last Login
customer.lastLogin
Date New Draft

SUGGESTED ACTIONS:

Customer actions
13 suggestions

Rules & Validation
Business rules triggered 5

Email Format Validation
RFC 5322 standard check

Phone Number Format
E.164 format required

Required Fields Check
Validate mandatory fields

Special Character Check
Remove invalid characters from fields

Null Value Handling
Flag or reject empty required fields

Data Enrichment
Fill missing data 3

Duplicate Detection
Merge duplicate records 3

AI Agents
AI copilot & automation 2

Order Transactions actions
8 suggestions

Product actions
4 suggestions

The result

KOMATSU[®]



1.

There is now one, **just one** person that maintains the full engine today. This used to be a full team.



2.

Komatsu have scaled to bring in all their product data



3.

Fabric (F128) is the central compute engine where other technologies use it for general computer

ROADS & INFRASTRUCTURE

80% of the tasks I would have handed to developers, I can now do myself in one minute.

FREDRIK KERSTAN



WEALTH MANAGEMENT

We chose CluedIn because we share a common vision of accelerating value creation through the responsible democratization of data.

PERNILLE OLLING
SENIOR VICE PRESIDENT

Nykredit



MEDIA AND ENTERTAINMENT

We chose CluedIn for its powerful AI capabilities and seamless Azure-native integrations to empower our business and technical teams with data that's ready for insights & AI-driven innovation.

FELIX BAKER
HEAD OF DATA SERVICES

SEGA



RETAIL

With CluedIn, we can process data as-is and adjust our schema adaptively.

LAURETTE CHANG-LENG
HEAD OF INTEGRATION

MITRE 10



RISK MANAGEMENT & INSURANCE

As a Graph-based system, we have experienced firsthand how easy it is to ingest and map data, and how simple it is to define our data model.

ROBERT BROWN
ENTERPRISE ARCHITECT

GUARDRISK



AVIATION

I hope you do more of these because we all have the exact same issues, and that comes out as soon as we get together in a room like this.

FABIO ASSMAN
DIRECTOR OF ANALYTICS & DATA GOVERNANCE

JET AVIATION



Microsoft

Marketplace Ratings

★★★★☆ 4.8 Star Rating



Gartner
peerinsights™

★★★★☆ 4.8 Star Rating

PLAINS

GUARDRISK



KANTAR

Nykredit

SEGA

MITRE 10



AMSIVE

Barentz

JET AVIATION



Urban Partners

ICG

Gallagher

INTERNATIONAL MISSION BOARD

ADX

EVERCORE

RIYADH AIR

MASC

Clear.Bank

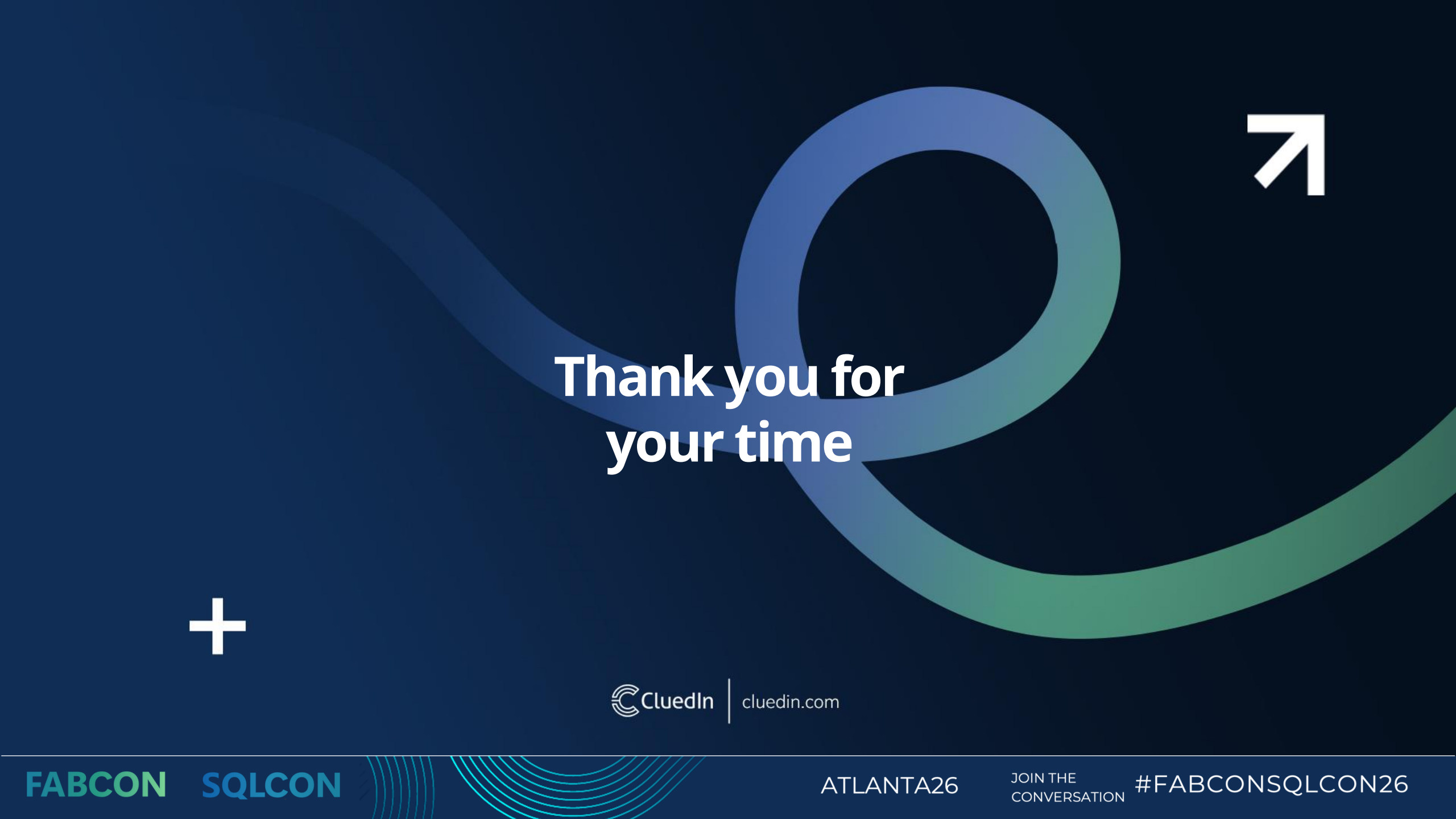
FABCON

SQLCON

ATLANTA26

JOIN THE
CONVERSATION

#FABCONSQLCON26



Thank you for
your time

 CluedIn | cluedin.com

Sound off.
The mic is all yours.
Influence the product roadmap.

Join the Fabric User Panel



Share your feedback directly with our
Fabric product group and researchers.

<https://aka.ms/JoinFabricUserPanel>

Join the SQL User Panel



Influence our SQL roadmap and ensure
it meets your real-life needs

<https://aka.ms/JoinSQLUserPanel>

How was the session?



Complete Session Surveys in
Whova for your chance to WIN
PRIZES!

