

# Your Data Is Fine. Your AI Isn't Good Enough. Seriously.



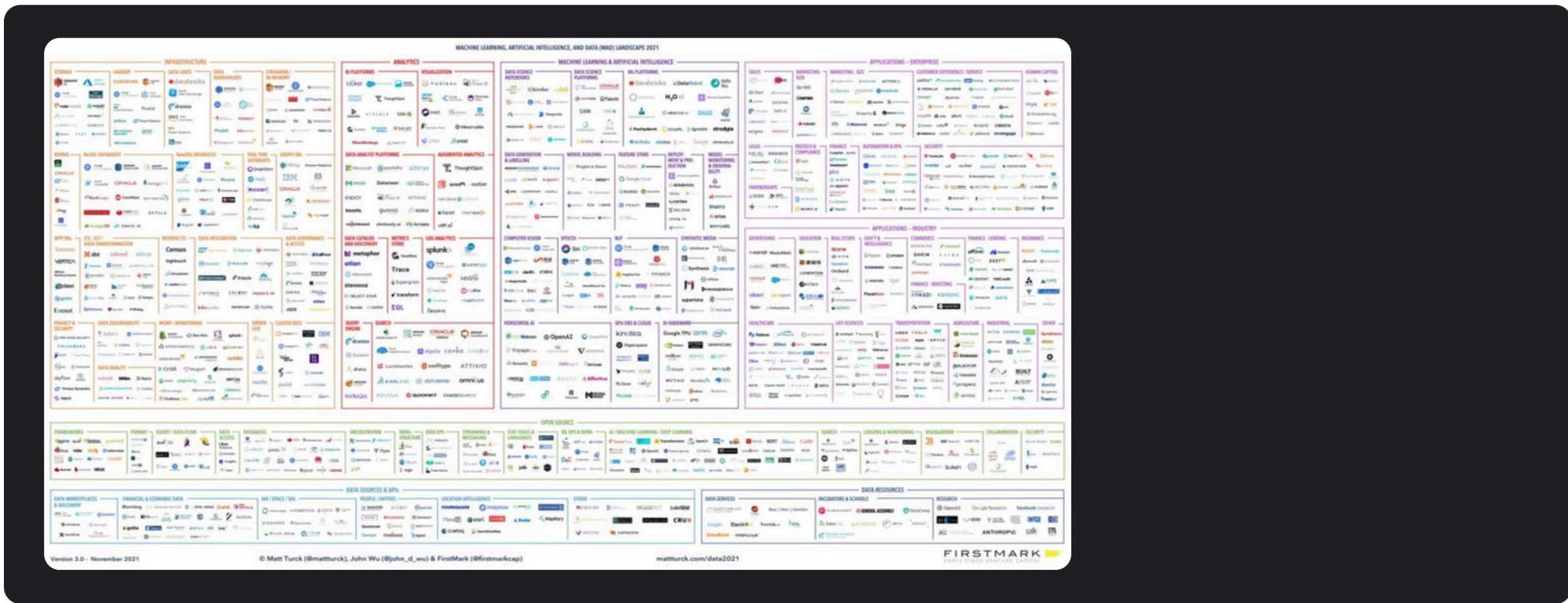
**Ethan Ding**  
CEO; Co-founder



**Rob Wisniewski**  
CTO, BXCI Blackstone



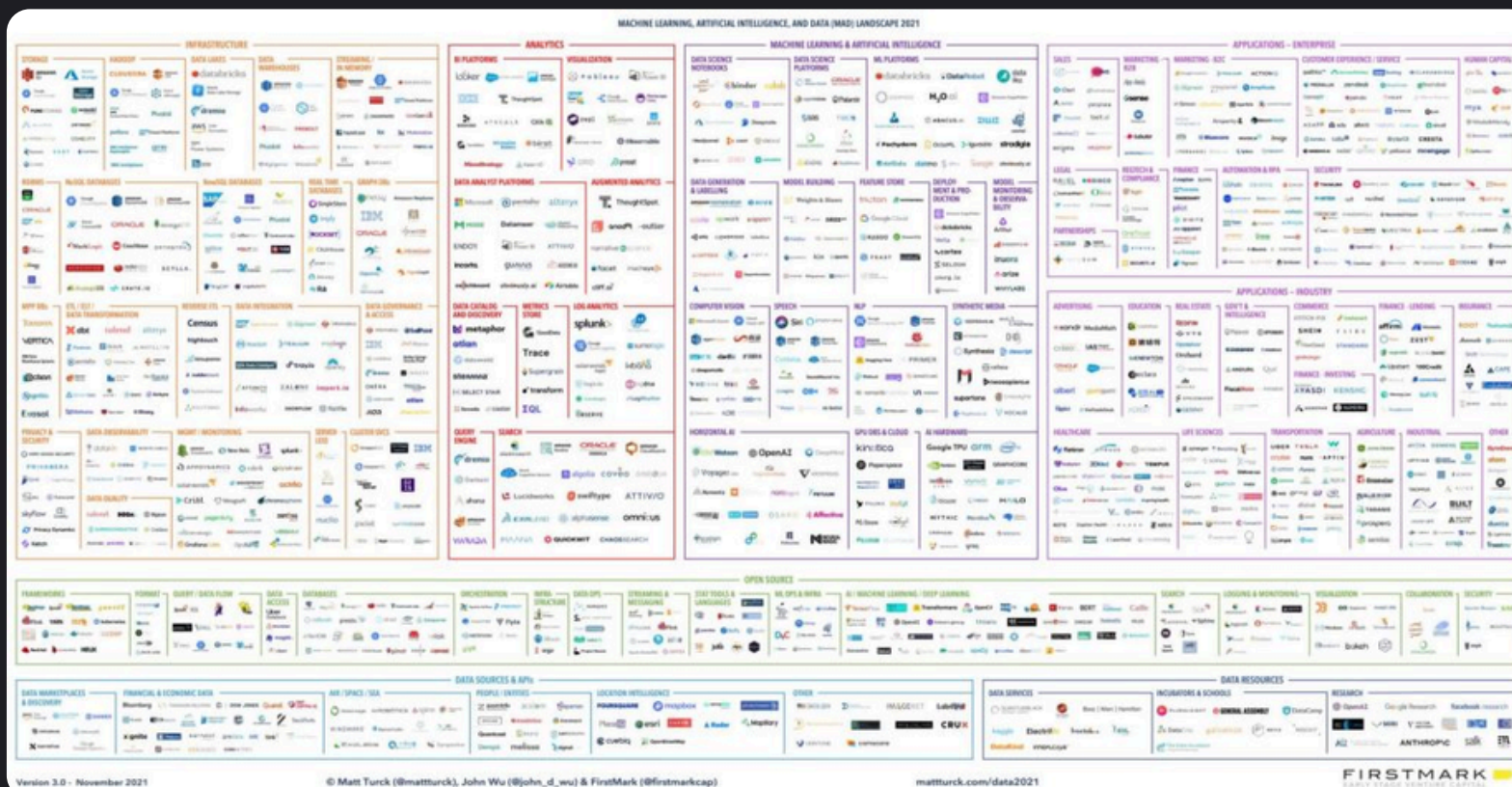
# This is the data stack you want to self service analytics on



These are the vendors who want to let you do “Chat with your Data”



# These are the vendors who want to let you do “Chat with your Data”



# These are the vendors who want to let you do “Chat with your Data”

The image displays a comprehensive 'MACHINE LEARNING, ARTIFICIAL INTELLIGENCE, AND DATA (MAD) LANDSCAPE 2021'. It is a dense grid of vendor logos organized into various categories such as Infrastructure, Analytics, Machine Learning & Artificial Intelligence, Applications - Enterprise, and Applications - Industry. A large white speech bubble is overlaid on the center of the grid, containing the text: "Just move your data here then clean it then your AI will be ready!". At the bottom of the grid, there is a footer with the following information: "Version 3.0 - November 2021", "© Matt Turck (@mattturck), John Wu (@john\_d\_wu) & FirstMark (@firstmarkcap)", "mattturck.com/data2021", and the "FIRSTMARK" logo.

# These are the vendors who want to let you do “Chat with your Data”



- Just move your data here
- Just move your data here and clean it
- Just move your data here and clean it with a team of consultants
- Just move your data here and clean it with a team of consultants for 5 years

... then it'll be ready





If your data is good enough → Your AI will be good!

Who We Are

AI-Driven Data Workloads

An Agentic Approach

TextQL In Action

Getting Started



~~If your data is good enough →~~

**If your AI is good enough →**

~~Your AI will be good!~~

**Your Data will be good  
enough**

But building an AI that can clean up the data of a F500 company is hard, they average...

**~150,000  
Tables**

across 2 Cloud  
Providers & 3 Data  
Warehouses

**~10,000  
Dashboards**

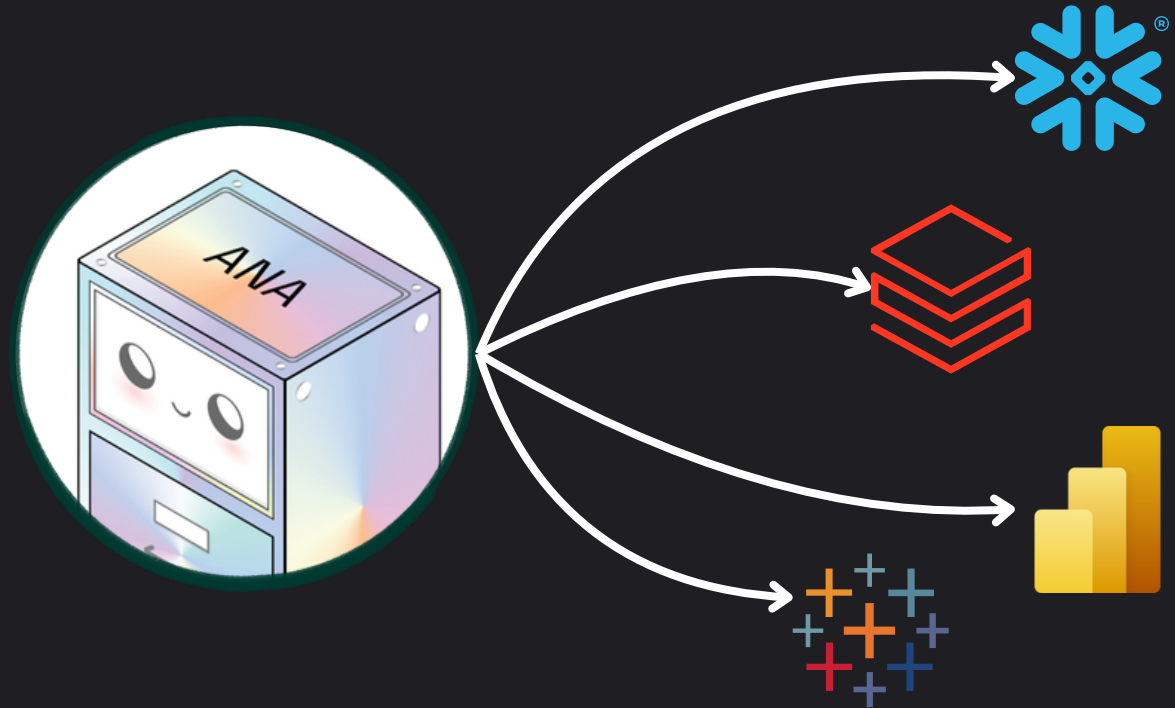
across 4 Business  
Intelligence Tools, On  
Prem & Cloud

**<1% of Data  
Catalogs  
Populated**

across Governance &  
Cataloging products

It'd require the AI to be able to go across on prem & in the cloud databases to...

- Query the data
  - From multiple environments
  - w/ multiple SQL syntaxes
- Load that data into a compute environment that can fit it
- Transform the data
- Eval against source of truth
- Store memory of what it retained



# TextQL Product Architecture

AGENTIC AI PLATFORM FOR ENTERPRISE DATA ANALYTICS

## USER INTERFACE & ACCESS LAYER

<b>Web Application</b> React / Streamlit	<b>Slack Integration</b> Bot + Webhooks	<b>API Endpoints</b> REST / GraphQL	<b>Playbook Scheduler</b> Cron-based Automation	<b>Dashboard Engine</b> Streamlit Runtime	<b>Console Portal</b> Billing & Admin
---	--	--	--	--	--

## AGENT ORCHESTRATION ENGINE

**Ana — Autonomous Data Agent**

MULTI-STEP REASONING    TOOL SELECTION    CONTEXT MANAGEMENT

<b>Text-to-SQL</b> Multi-model	<b>Python Sandbox</b> Isolated Execution	<b>Web Search</b> Live Search	<b>Report Builder</b> Structured Output	<b>API Orchestrator</b> 3rd Party Integrations	<b>Context Library</b> Knowledge Layer	<b>Viz Engine</b> Charts
-----------------------------------	---	----------------------------------	--	---	---	-----------------------------

## DATA INFRASTRUCTURE & CONNECTORS

UNIVERSAL CONNECTOR FRAMEWORK

<b>Snowflake</b>	<b>BigQuery</b>	<b>Redshift</b>	<b>Aurora / Postgres</b>	<b>Databricks</b>	<b>Athena</b>	<b>Azure Synapse</b>	<b>MySQL</b>	<b>SAP HANA</b>	<b>Supabase</b>	<b>ClickHouse</b>
------------------	-----------------	-----------------	--------------------------	-------------------	---------------	----------------------	--------------	-----------------	-----------------	-------------------

INTERNAL DATA PLATFORM

<b>Production DB</b> Aurora PostgresQL Schema Migrations Live Query / Changelog	<b>Console DB</b> Aurora PostgresQL ACU + Billing Tenant Config	<b>GTM Warehouse</b> Segment Attribution Tracking
<b>Changelog Tables</b> Logtail + Streaming	<b>Metering Engine</b> ACU Computation	<b>Webhooks ETL</b> Lambda Pipeline

EXTERNAL SAAS INTEGRATIONS

<b>Attio CRM</b> Pipeline / Contacts	<b>Stripe</b> Billing	<b>Grain</b> Recordings	<b>Loops</b> Email Marketing
<b>Notion</b> Docs / Wiki	<b>Linear</b> Issue Tracking	<b>Posthog</b> Product Analytics	<b>Slack</b> Communication

## INTELLIGENCE LAYER

<b>Ontology Engine</b> Schema Discovery Semantic Mapping	<b>Metric Definitions</b> Expressed Formulas Business Metrics	<b>Context Repository</b> Context Injection Knowledge Graph
<b>LLM Router</b> Model Selection Fallback Chains Latency Routing	<b>ACU Metering</b> Token Counting Query Complexity Cost Attribution	<b>Pricing Engine</b> ACU Rates Billing Thresholds Overages

## DELIVERY & OPERATIONS

<b>Report Generator</b> Scheduled Reports PDF / HTML Output	<b>Slack Delivery</b> Channel Posts Scheduled Options	<b>Email Delivery</b> Scheduled Reports Branded HTML
<b>Playbook Engine</b> Cron Jobs Step Flows REST / HTTP	<b>Dashboard Runtime</b> Streamlit Sender Live Refresh Embed Support	<b>Monitoring</b> Error Alerts SLA Tracking Outgoing SPM

## SECURITY, GOVERNANCE & INFRASTRUCTURE

<b>IAM &amp; RBAC</b> Role Management Fine-grained Access	<b>Secrets Vault</b> AWS Secrets Mgr KMS Encryption	<b>Data Isolation</b> Tenant Separation VPC Isolation	<b>AWS Infrastructure</b> EC2 / Lambda CloudFront / S3	<b>Audit Trail</b> Immutable Logging Query History	<b>Multi-tenant</b> Org Isolation Custom Config Shared Defs	<b>CI/CD</b> GitHub Actions Preview Envs	<b>SOC 2 / HIPAA</b> Compliance Certifications
---	---	---	--	--	--	--	--

But if it can do all those things, then...



# Evaluating what makes for an AI that can query Enterprise Data



**Rob Wisniewski**  
CTO, BXCI Blackstone

---

Blackstone

We liked hearing that this would work on messy data, so we wanted to run the “easy” test of...



## SQL Server to Snowflake Migrations

**Pulling unified analytics across Blackstone Credit & Insurance Snowflake - no data dictionary, no semantic layer, no config. Just connect.**

50+  
siloed Credit &  
Insurance teams

1,000s  
tables, zero shared  
catalog

~12  
competing metric  
definition

0  
agreed upon  
semantic layer

18 mo  
previously spent  
trying to solve this

- "Net exposure" means something different in every team schema - different tables, different denominators
- Credit and Insurance teams share no common metric definitions - business logic lives in analyst heads
- Cross-team queries routinely fail. No one has successfully queried across both sides at once.

Could AI Make this easier?

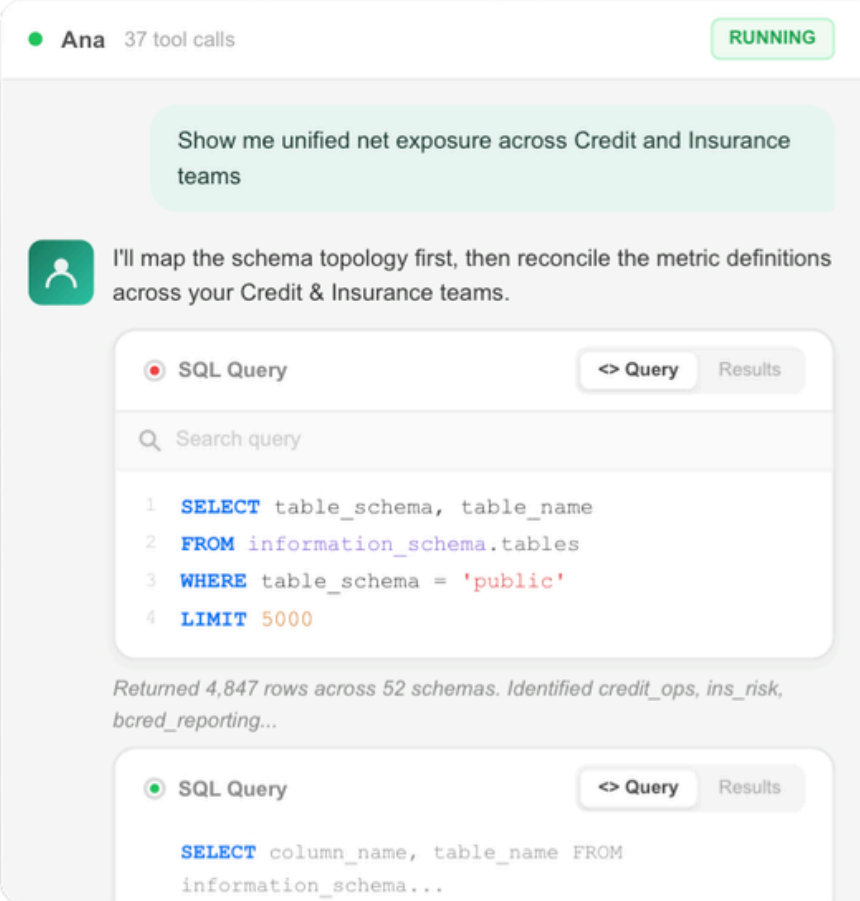
# ...Ana taught herself the warehouse, then reconciled 18 months of unsolved complexity

1. Self-directed schema discovery
  - Ana mapped 52 schemas and flagged metric divergences before writing a single query
2. Cross-team reconciliation
  - Unified the same labels across different tables and denominators
3. Zero configuration
  - “We spent 18 months trying to get these teams to agree on a number. Ana just did it.”

<4 minutes  
time to first cross-team  
insight

0  
config files or set up  
required

37  
autonomous tool calls



Ana 37 tool calls RUNNING

Show me unified net exposure across Credit and Insurance teams

I'll map the schema topology first, then reconcile the metric definitions across your Credit & Insurance teams.

SQL Query <> Query Results

Search query

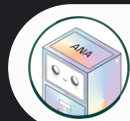
```
1 SELECT table_schema, table_name
2 FROM information_schema.tables
3 WHERE table_schema = 'public'
4 LIMIT 5000
```

Returned 4,847 rows across 52 schemas. Identified credit\_ops, ins\_risk, bcred\_reporting...

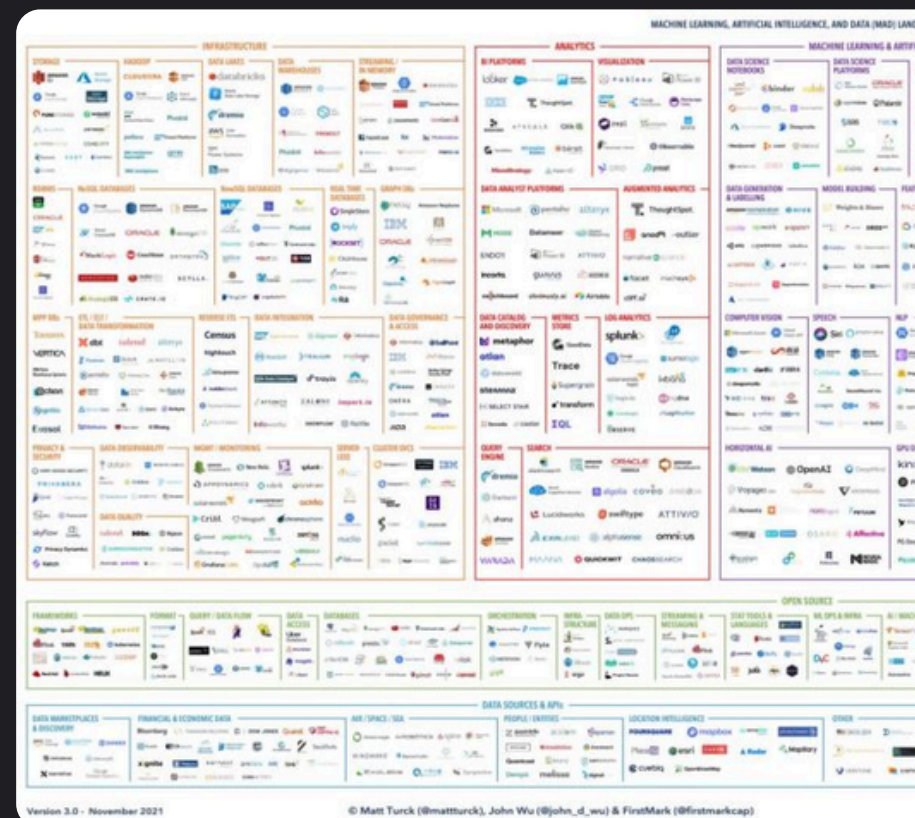
SQL Query <> Query Results

```
SELECT column_name, table_name FROM
information_schema...
```

# What can an Agent that can query messy & complex data do for the TCO of your data stack



Hey, can you figure out the messiest parts of my data... and clean it up?



# What is everyone missing about the world where AI can prepare your data for you?

# Make us Prove it

## Try today:

1. Go to [app.textql.com](https://app.textql.com)
2. Connect data
3. Ask Ana “can you prepare my data for analysis”

## Deploy Pilot:

1. Go to [textql.com/request-data](https://textql.com/request-data)
2. Enter GARTNER-TALK in “how you heard of us”
3. Get \$25-150K in POC funding for qualified enterprises