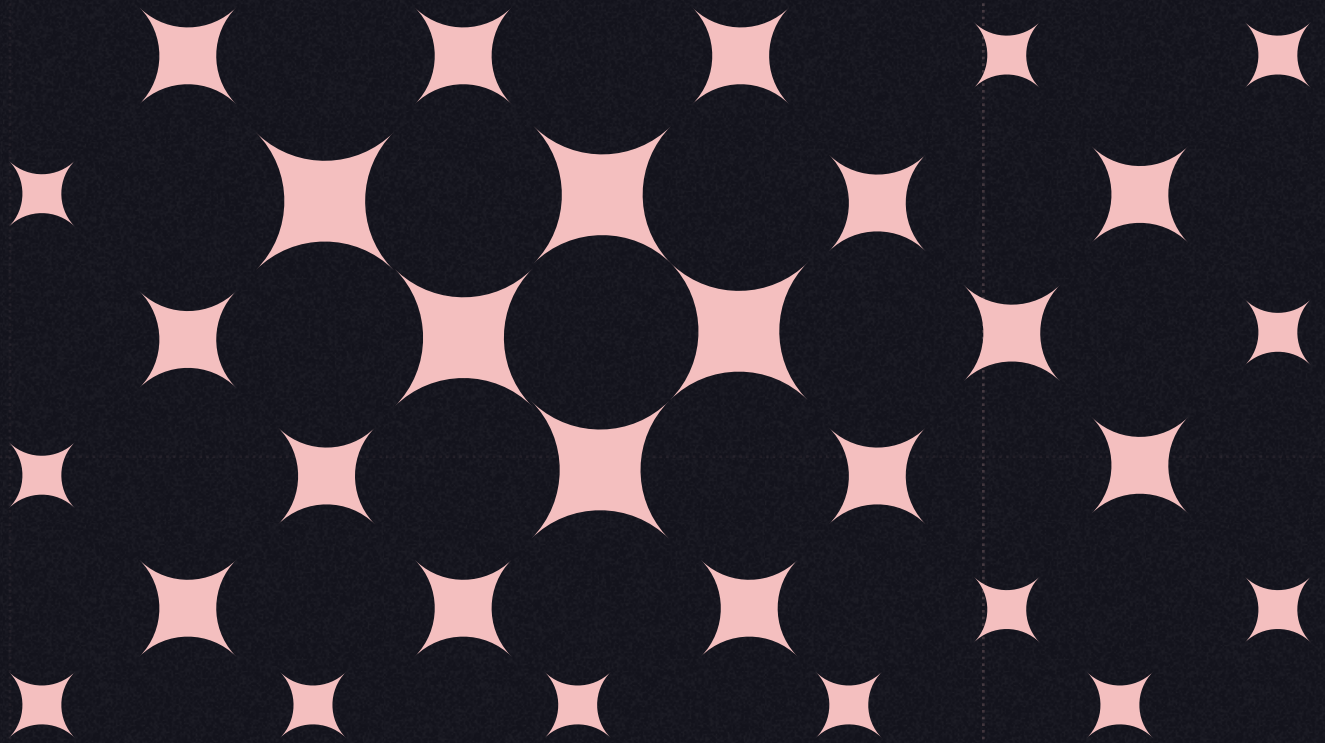
HEX

www.hex.tech

A practical roadmap for understanding and implementing context engineering to power self-service agentic analytics

The Data Leader's Playbook for Agentic Analytics



Make everyone a data person

Dashboards were never the destination

Your data team is drowning in requests. Every day brings another Slack message: "Can you pull this metric?" "Quick question about last quarter's..." The promise of self-service analytics has been just that — a promise.

AI is changing that equation, but it requires a fundamental shift. Data teams need new skills, new workflows, and a willingness to challenge the assumption that dashboards are the pinnacle of self-service. This is happening fast. Teams that don't adapt will remain the bottleneck indefinitely.

What users actually want isn't more dashboards — it's insights on demand. Asking questions in plain language, getting trustworthy answers from governed sources, and confidently making decisions.

Self-service makes leaders nervous because previous attempts have failed. The difference now: **AI makes it possible, but only when paired with proper infrastructure.** You need two things: context that grounds the agent in your reality, and observability to systematically improve it.

This is Hex's thesis: you need a platform where agents and humans collaborate in one connected workflow — data teams accelerate their work while business users get governed self-service insights.

Not a chatbot bolted onto your BI tool. A system designed for this from the ground up.

After reading this guide, you'll walk away with:

- **Why context matters and clarity on where you should start**
- **A 2 week playbook to see value with Threads, Hex's conversational analytics feature**

Teams that succeed master this loop:

1. **Pick your use case**
2. **Curate your context**
3. **Understand why answers were wrong, and how to fix**
4. **Run a pilot with 3-5 non-data team members**
5. **Iterate and improve**

The outcome? Your data team delivers value faster. Instead of a bottleneck, you become the architects of a self-service system that scales — spending less time answering "what happened" and more time influencing "what should we do."

Context — your foundation for success

In a recent survey, 77% of data leaders said they were excited about the possibilities of AI, but only 3% made it a top priority, citing data quality and reliability challenges as a major obstacle.

There are thousands of “AI-powered analytics” offerings out there, representing a massive opportunity to make data-powered insights easier and faster. The risk? Agents that hallucinate, teams that lose trust, and pilots that stall.

That’s because in many existing technologies, context is an afterthought. But if you’re not thinking about context, you’re evaluating these tools wrong.

Traditional BI RFPs focus on dashboards, query performance, and user seats. AI-powered analytics evaluations need to focus on a different set of critical questions:

- **How do you observe what it's doing?**
- **What context does the agent use?**
- **Can you unilaterally improve it?**

Context engineering is becoming table stakes across all AI tooling. Investing here builds organizational capability that extends to every AI initiative you run.

Our belief: context isn't a tax, it's an asset.

The warehouse descriptions, documentation, and governance your team already maintains? That's context. We're just making it work for both AI and humans. Each improvement compounds — better agent responses lead to more usage, which surfaces more opportunities to refine. Hex’s Context Studio allows you to see and manage usage patterns, context assets, and agent performance all in one place.

Data teams that scale AI aren't the ones with the fanciest models — they're the ones who observe, evaluate, and improve context.

Context categories

Endorsed statuses

Your warehouse guardrails

Direct agents to your best data. Endorsed statuses act as quality signals, telling agents which tables to prioritize.

Learn more ⟶

Warehouse descriptions

The foundational context

These are your standard table and column descriptions — basic hygiene that pays dividends with AI.

Learn more ⟶

Workspace guides

Teaching the agent your business

Add the tribal knowledge that doesn't live in your database: business terminology, metric selection criteria, data quirks, seasonal patterns, and workflow preferences. This helps the agent think like your data team.

Learn more ⟶

Semantic models

The rigid rules for agents

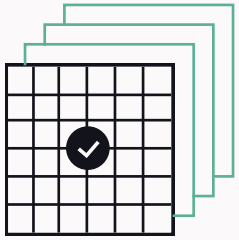
When metrics must be 100% correct every time, semantic models enforce table relationships, measure calculations, and available dimensions.

Learn more ⟶

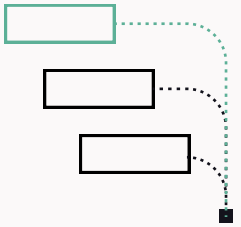
Where to start with context engineering

If you're new to agents & context engineering, spend 30 minutes endorsing your trusted tables in Hex and confirming they have good warehouse descriptions. Our agents prioritize endorsed tables and semantic models; endorsements are the fastest way to improve responses.

Your order should be:

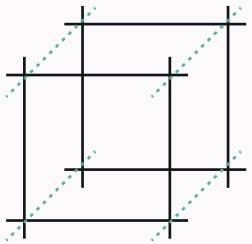


1 **Endorsing key tables and/or semantic models, exclude tables/ schemas that shouldn't be used**



2 **Adding warehouse descriptions on tables or columns that are important, and have been endorsed**

3 **Creating a workspace guide with 5–10 rules**



4 **Adding semantic models**

Speed + accuracy

> *We had a product manager who needed to understand Chrome for Android usage before sprint planning. Being 3% off didn't matter. Getting the answer in 5 minutes instead of 2 days absolutely did."*

Hannah Burak, Panda Doc

Speed matters — but only when it's paired with a system that continuously increases accuracy. No AI system is perfectly accurate on day one, and waiting for perfection is the fastest way to stall progress.

Measuring improvement over time matters more than single-answer accuracy. Can the agent read new context?

Can you quickly refine and see improvement? That's the acceleration pedal — not obsessing over precision on individual queries.

Consider what business users do today. They either make gut calls or they don't ask at all because the queue is a black hole — two weeks minimum, if it gets prioritized.

Here's what accuracy actually means in an agentic analytics system:

Trusted data is always prioritized	Endorsements and guardrails prevent agents from touching ungoverned sources.
High-stakes questions route through governed logic	Revenue, ARR, active users — these come from semantic models that enforce exactness.
Low-stakes exploratory questions get speed first, accuracy second	This mirrors what business teams already do today, but with far more visibility.
Every incorrect answer improves the system	Fixing context once improves it for all future questions — not just the single thread.

Now they can ask the agent, get a directional answer, validate it with the data team if needed, and move forward. The goal is not to relax accuracy standards. The goal is to apply the right level of accuracy to the right type of question and improve the baseline continuously.

If it's not accuracy, what should I be evaluating?

Accuracy matters, but evaluating an AI analytics system solely on "did it get this one question right?" misses the real point.

In a governed, context-driven system, accuracy is the outcome of the governance layer, not the metric you use to determine fit. That's what separates toy chatbots from enterprise systems.

Questions where directionally-right answers are ok
"What are the most common screen sizes for Chrome users?"

Questions where the answers must be exactly right every time
"What is our monthly sales revenue for each month from last quarter?"

Exploration and discovery - there might not be a right answer
"Why didn't anyone read the blog we posted last week?"

You should evaluate on four dimensions:

Answer Quality & Relevance

- Can the data team rate accuracy consistently?
- Does the agent consistently use the right governed metric?
- Do end-users find answers relevant, not just technically correct?

Collaboration & Error Resolution

- How easily can users share conversations with the data team for validation?
- When agents get it wrong, does fixing it improve future responses or just patch one answer?

Observability & Systematic Improvement

- Can you see what users are asking and how agents are responding?
- Does the tool allow you to quickly improve context?

Ecosystem Integration

- Does it work with your existing data stack and governance?
- Do context improvements benefit other agents, or stay siloed to just one?

The strategic shift

The reality

Data teams have been stuck being the service desk because self-service tools either gave users too much rope (and wrong answers) or not enough flexibility (and couldn't answer real questions).

AI agents — when properly grounded in context — thread that needle. They're powerful enough to handle complex analytical work, but constrained enough to stay trustworthy.

Your job as a data leader is to build that context infrastructure. To curate what's trustworthy, document what's nuanced, and model what's complex. To turn your institutional knowledge into machine-readable instructions.

Done right, you're no longer just serving requests — you're building the foundation for every data decision that follows.

The data team that embraces this shift isn't just adopting new tools. They're reclaiming their time to work on the problems that actually matter — understanding what questions are important to the business, weighing in on strategy, and building the data infrastructure that enables better decisions at every level.

Ready to start?

1. **Turn on Threads for your data team**
2. **Pick your first use case**
3. **Curate your context**
4. **Understand why answers were wrong, and how to fix**
5. **Run a pilot with 5 non-data team members**
6. **Iterate and improve**

The future of data work isn't writing more SQL. It's building the context that lets everyone answer their own questions — accurately, quickly, and at scale.

From bottleneck to architect

Scaling with Threads

New users who engage with Threads are **4-5x more likely** to remain active in Hex over subsequent weeks.

The key to scaling data to your entire organization is great self-service insights, not more dashboards. Threads is Hex's way of providing it — through an intuitive, conversational system that can deliver deep answers via chat.

Threads represents a new approach to conversational analytics and business intelligence that focuses less on reports and more on enabling business stakeholders to ask their own questions. This, in turn, allows the data team to focus on more scalable efforts than answering one-off business questions.

The rest of this guide focuses on a 2-week evaluation of Threads, leveraging all of the criteria we outlined previously.

Your first two weeks with Threads

If you're a data leader, use this evaluation timeline to assess the real impact of conversational analytics. You'll get a sense of: 1) how well your team can add/manage context, and 2) how well that context improves responses for regular end users.

Week 1

Prove value fast

Outcome

You establish your base context and understand wins/gaps.

- **Select your first use case, something like "understanding user engagement KPIs"**
  - Select a use case that your team get's "quick question" about often or the business has already aligned on it's definition with trusted data.
  - Align on the 5-7 questions you want to test with.
- **Turn on Threads for the data team only**
  - Connect Threads to Slack to amplify questions and responses.
- **Endorse your "golden" tables/semantic models, including those for your use case**
- **Add/upgrade the top 10 descriptions that matter**
  - Use descriptions that help the agent understand what kind of data exists in a table or column. Think, "would a jr. analyst read this description and understand what this data is?"
  - Make sure relevant columns have descriptions indicating common join conditions and whether they are 1-to-1 or 1-to-many.
- **Create a 1-page workspace guide (5–10 rules)**
  - You can upload existing documentation or ideas into ChatGPT/ Claude and get it to write your workspace guide.
- **Review responses, improve context where necessary and collect wins/gaps**

Week 2

Make it repeatable

Outcome

Refine your context and expand its scope to answer business users' questions.

Pilot with 5–10 motivated business users; collect structured feedback

Review their feedback and questions in the Slack channel or space you're gathering this information:

What questions worked great?	What questions didn't work?
What context was missing?	Where did they get stuck?

Review engagement and responses in the Context Studio.

- Set baseline adoption metrics and review in Hex
- Encourage users to give a thumbs up/down to responses for faster triage
- Dig into Threads that gave incorrect responses to identify what context is missing

Expand your context based on their questions and agent responses.

Optional

Add a lightweight semantic model where precision matters.

Hex's agents

Software engineers have been the primary beneficiaries of AI, helping them write code and create entire apps with a single prompt. Data teams can now leverage that same horsepower with Hex's agents, observing their performance in the Context Studio.

Improving a single agent benefits all — the context and data products created become sources of truth for end users via chat. **In Hex, you're not choosing between doing your regular work and improving the agent; your regular work is improving the agent.**

The Notebook Agent

Your data team's accelerator

The Notebook Agent accelerates the deep analytics work of your analysts, analytics engineers, and data scientists.

The outcome

- Complex SQL queries & Python scripts written 10x faster
- Immediate insights and recommended next steps in your analysis
- Interactive data apps developed faster, with natural language

The Modeling Agent

Your context curator

The Modeling Agent creates and governs semantic models, rapidly accelerating their development by synthesizing existing analyses from Hex projects into codified logic.

The outcome

- Semantic models leverage trusted sources, because they're built from existing Hex projects
- Rapid iteration and validation with AI-powered suggestions
- Models build on best practices and are backed by Hex documentation

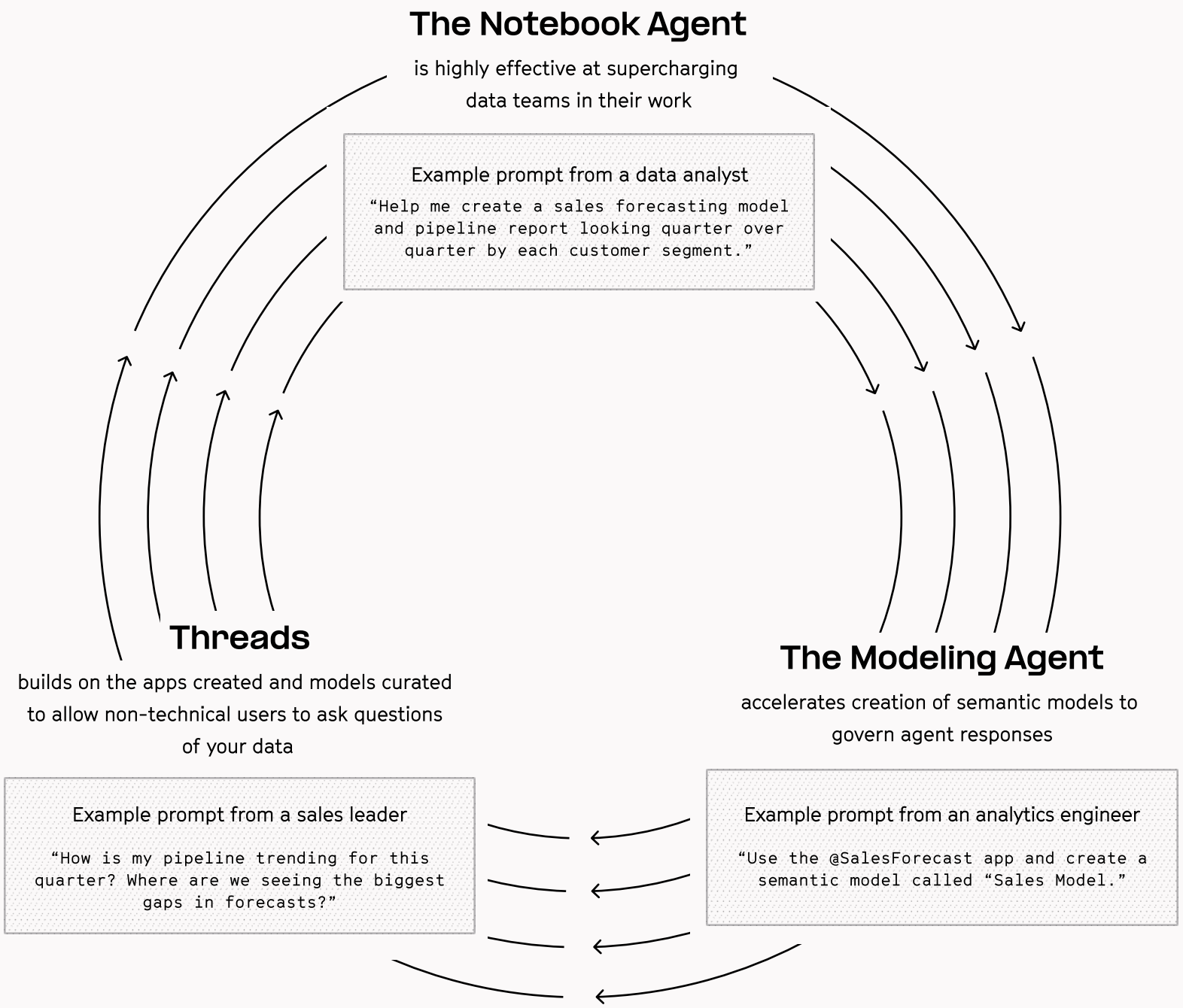
Threads

Conversational analytics for your business users

Threads enables conversational self-serve for your product managers, executives, sales ops, and anyone who needs answers from data but doesn't write code.

The outcome

- Faster decision-making across the org
- Reduced data team queue times for simple questions
- Conversational analytics that's governed by your context


The Notebook Agent
is highly effective at supercharging data teams in their work
Example prompt from a data analyst
“Help me create a sales forecasting model and pipeline report looking quarter over quarter by each customer segment.”
Threads
builds on the apps created and models curated to allow non-technical users to ask questions of your data
The Modeling Agent
accelerates creation of semantic models to govern agent responses
Example prompt from a sales leader
“How is my pipeline trending for this quarter? Where are we seeing the biggest gaps in forecasts?”
Example prompt from an analytics engineer
“Use the @SalesForecast app and create a semantic model called “Sales Model.”

HEX

Hex brings the magic of AI to data analytics, for everyone. It's the first platform for truly agentic analysis – where users of all technicality levels can collaborate with each other and AI to better understand their world, find insights, and make better decisions.