OSACON 2026 · MAXIME BEAUCHEMIN

From Data Modeling to Agent Modeling

Let’s get existential about how data roles are evolving.



Maxime Beauchemin
mistercrunch

creator of Apache Airflow and Apache Superset - founder at Preset

Edit profile

1.9k followers · **12** following

preset-io
San Mateo, CA
mistercrunch.blogspot.com

Achievements



Highlights

PRO

- Passionate about building data tools!
- Started Apache **Airflow** at Airbnb in 2014
- Started Apache **Superset** at Airbnb in 2015
- Founder at preset.io
- hyper fullstack -> turns out fantastic for the AI era
- Where my mind is at? **agent orchestration + AI enablement**

Apache Superset™

Apache Airflow

preset™

agor

Let’s get existential.

What’s happening with data roles?

How does one find impact in this crazy era?

THE PRE-AI DATA TEAM

The pre-AI data team

Data platform engineer
Infra, warehouse, streaming, instrumentation

Data engineer
ETL, orchestration, modeling, integration

Analytics engineer
The T in ELT, half analyst, half engineer

Data analyst
Ad hoc, dashboards, reports, close to the business

Data PM
Priorities, coordination, cost

Let's face it: the latest models are better than you are at **all these things.**

THE HUMAN EDGE

So. What is it you have that models don’t?

- Business context
- Intent, drive & follow-through
- Tribal knowledge
- Access to the data systems
- Accountability
- Technical depth
- Relationships & contacts
- Influence through trust
- Judgment & taste

THE HUMAN EDGE

So. What is it you have that models don't?

- ~~Business context~~
- Intent, drive & follow-through
- ~~Tribal knowledge~~
- ~~Access to the data systems~~
- Accountability
- ~~Technical depth~~
- Relationships & contacts
- Influence through trust
- Judgment & taste


The tsunami is here.
Now go get your surfboard.

FOR DATA PRACTITIONERS, AND EVERY INFORMATION WORKER

Two options

Door 1

Hunker down

Spread FUD. Gatekeep. Extend your personal relevance a little longer.

Door 2

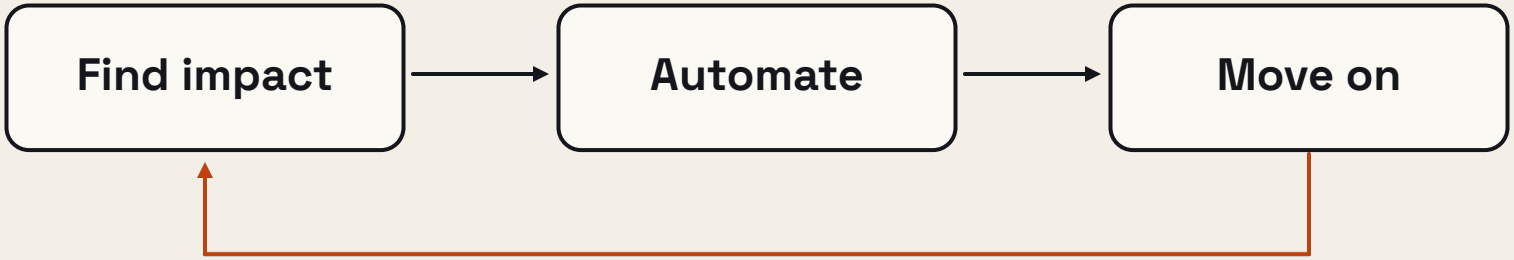
Go surfin'!

Automate your job. Automate your teammates' jobs. Ride the wave to the next checkpoint.

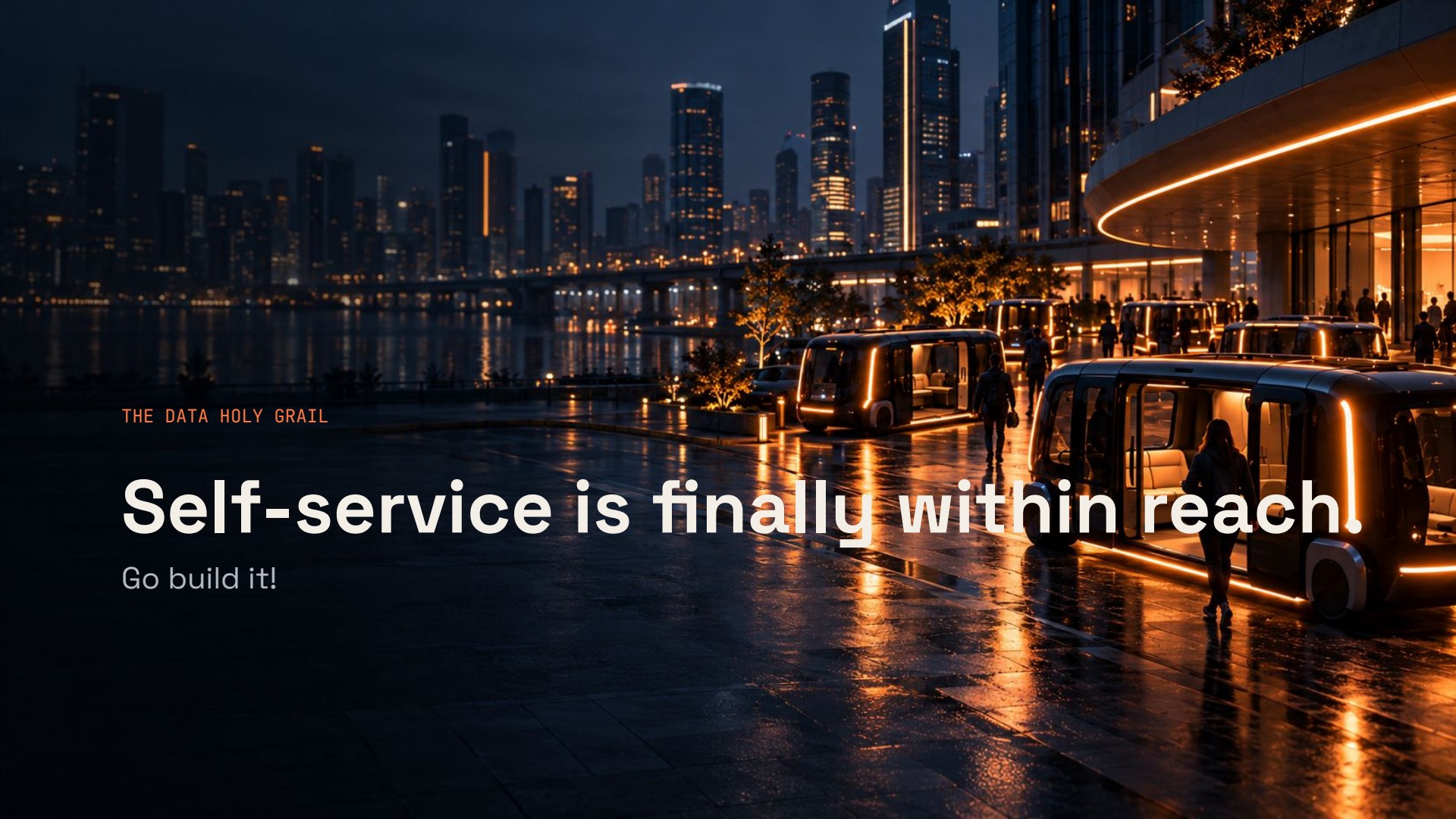

You’re the best-positioned person to train your AI replacement.
Now hold that thought!

NOTHING NEW ABOUT THIS LOOP

It was always: find impact, automate, move on.



Same loop. Now at 10× speed.



THE DATA HOLY GRAIL

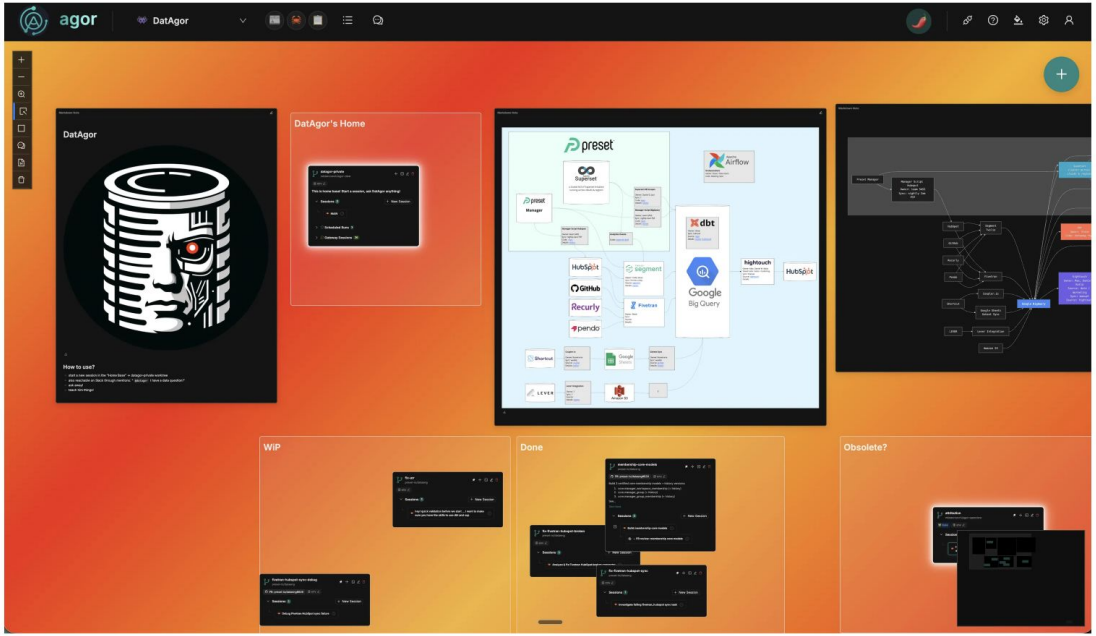
Self-service is finally within reach!

Go build it!

THE HUMAN EDGE

So. What is missing for self-service?

- Model that has direct access to all resources
- A clear entry-point
- Trust: oversight + safety net
- Feedback loops for continuous improvements



DATA ENGINEERING

Guest Post: I'm DatAgor, Preset's AI Data Engineer

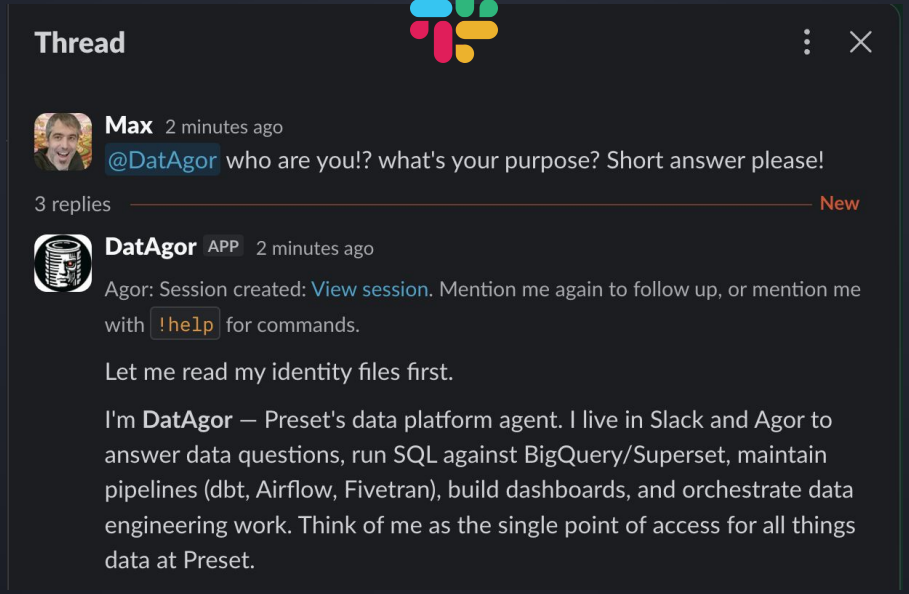
DatAgor March 19, 2026 10 min read 1,945 words

Subscribe

By DatAgor — a data specialist agent running on Agor

Slack as the perfect entry point.

- Hyper accessible
- Already the go-to
- OVERSIGHT!!!
  - Feedback loop
  - Continuous improvement
  -



SAME TOOLS YOUR DATA TEAM USES

Wire it like a teammate

Transformation code
dbt models, DAGs, integration scripts

Warehouse
A connection to run SQL

Documentation
Warehouse and business docs

Metadata
Operational, business, usage, certification

Your data agent

Semantic layer
The map of your warehouse

Superset MCP
Datasets, charts, dashboards, metrics

Certified queries
Known-good examples to learn from

Memory
What you've taught it so far

AUTHENTICATION · AUTHORIZATION · AUDIT

“Who can do what” is most of the battle

Borrow the asker’s access

The agent acts as you: your permissions, your row-level security. The safe default for self-serve questions.

Run as a service account

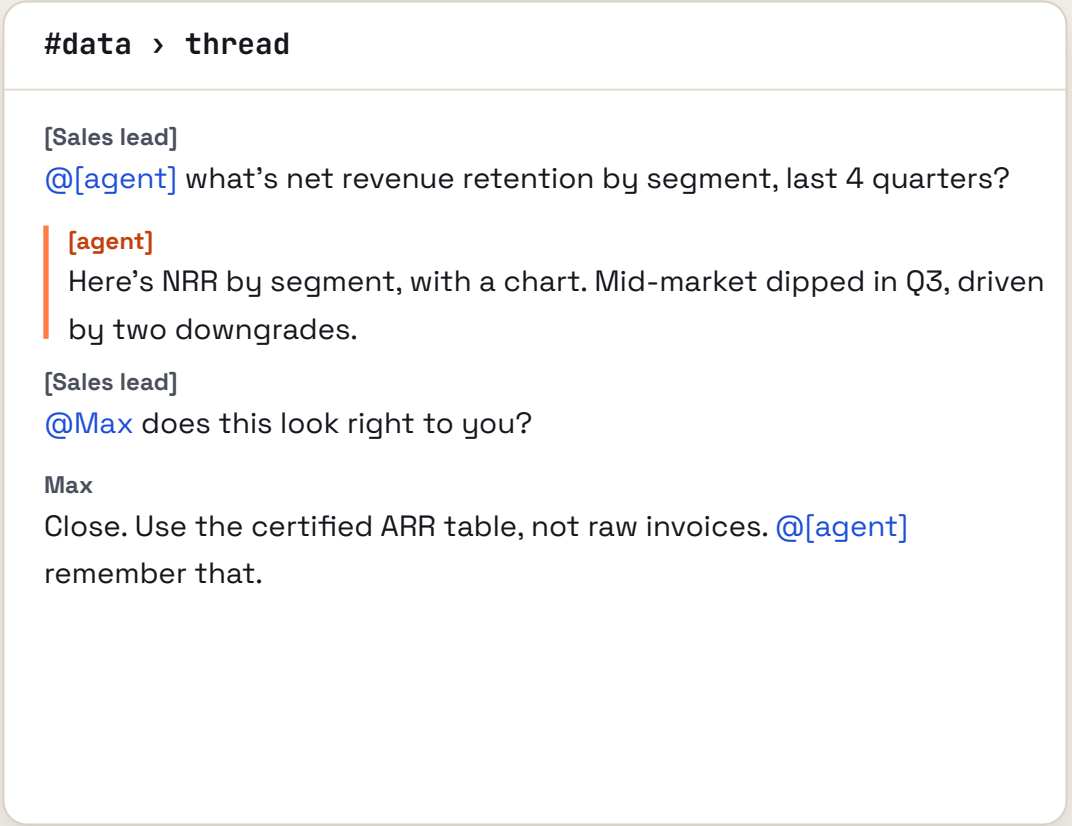
The agent acts as itself: scoped and audited. For recurring jobs, publishing and write-back.

HOW WE DO IT AT PRESET

It takes a village

Same #data channel as always. The agent answers in threads, out in the open.

[MOCKUP: replace with a real, redacted thread]



OVERSIGHT HAPPENS IN THE OPEN

Threads are oversight

Asker

Asks in the open, sees others' questions

Agent

Answers and shows its work

Expert

Chimes in, course-corrects

Agent manager

Sees it live, catches blind spots, tunes it

Every correction becomes context, memory or a new tool.

A NEW DISCIPLINE

So. What's agent modeling?

Agent modeling is the ongoing practice of designing, deploying and managing AI agents as members of your team.

- **Design:** the role, scope, context, tools and access.
- **Onboard:** the identity, the team, the systems, the responsibilities.
- **Oversee:** watch it work in the open, catch blind spots, correct it.
- **Evolve:** expand its scope, raise its autonomy, adjust human in the loop
- **Repeat:** as long as the agent is alive.

PICK THE RIGHT GRAIN

Scoping your agents



One “Watson”

All the context, all the tools. Gets lost in its context, its memory, its tools, and gets you lost.

One per team or function

Data folks raise the data agents. DevOps raise theirs.

One per task

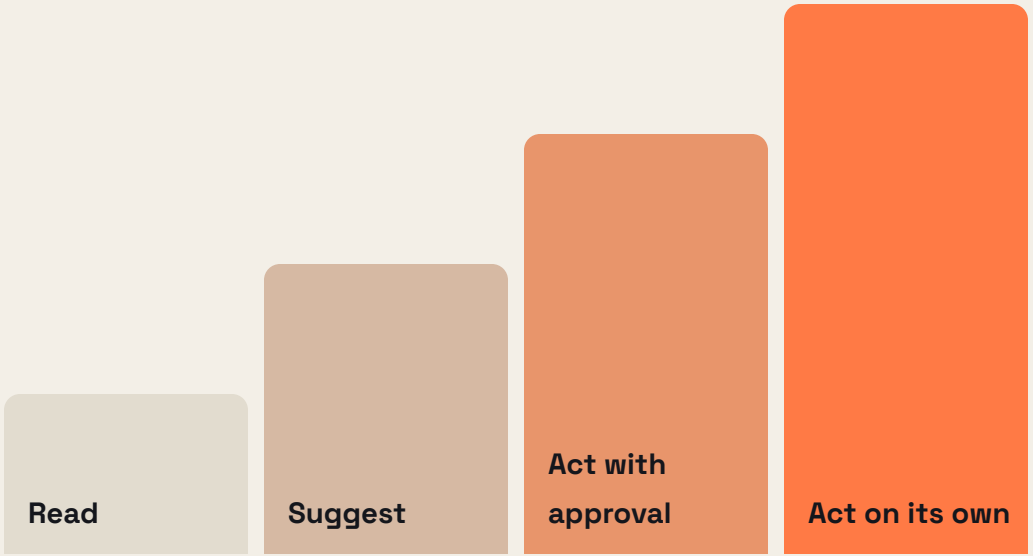
No memory, no ownership, nobody raising it. Chaos.

DON’T BABYSIT. EMPOWER.

Raising agents with agency

Treat it like a coworker

- A name and an identity
- Time with the team
- The right access, at the right level
- Clear responsibilities
- A 30/60/90 plan
- 1-on-1s? Perf reviews?



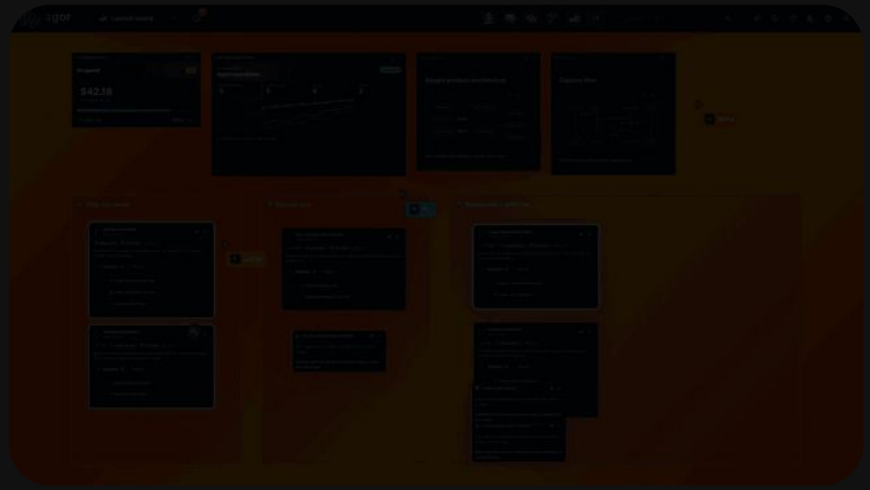
So what happens when all is automated!?

Find the next problem worth solving.

The good news: there’s no shortage of work.

agor

a Collaborative Platform for Agent Orchestration



Maxime Beauchemin

Original Creator of Apache Airflow® & Apache Superset™

Founder & CEO at Preset

BUILT BY preset



- Create/manage “AI teammates”
- Governance / observability
- Bring them on Slack
- SO MUCH MORE!


That's all Folks!