AI in BI:

The Path to Full Self-Driving Analytics

May 1st 2025

Maxime Beauchemin
mistercrunch

creator of Apache Airflow and Apache Superset - founder at Preset

Edit profile

1k followers · **11** following · ☆ **139**

preset-io
San Mateo, CA
maximebeauchemin@gmail.com
mistercrunch.blogspot.com

Organizations

- Passionate about building data tools!
- Started Apache **Airflow** at Airbnb in 2014
- Started Apache **Superset** at Airbnb in 2015
- Started **Preset** - The Apache Superset company in 2019

Apache Superset™

Apache Airflow

preset™

Some context

Qualified?

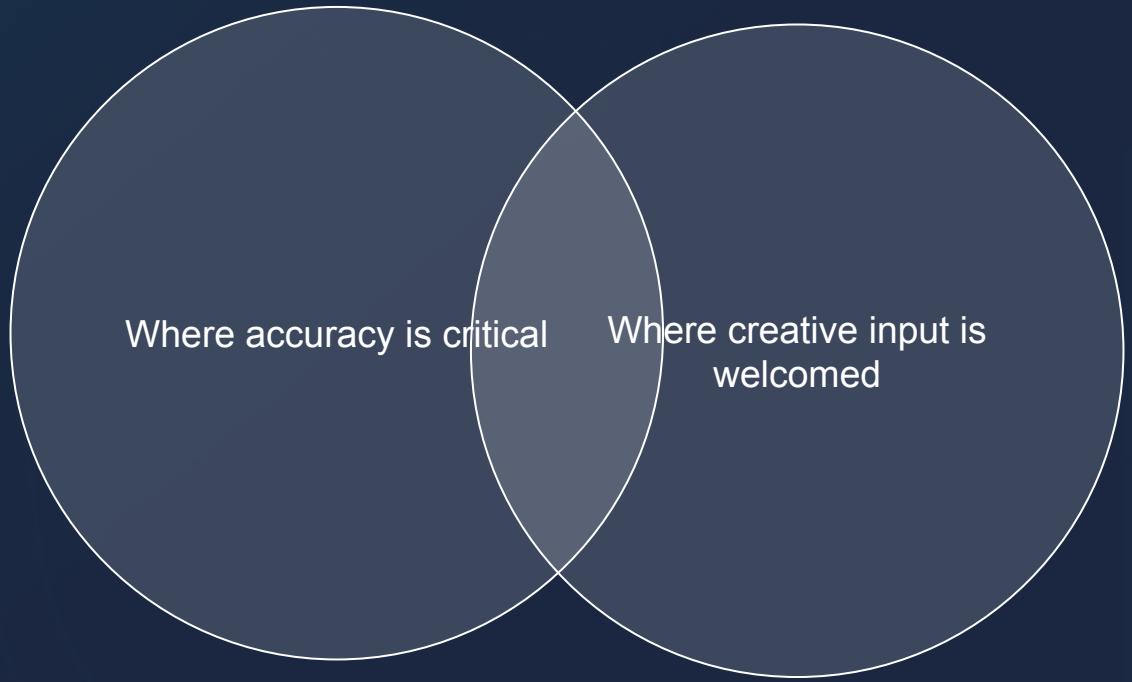
- Avid AI user since GPT 3.5
- Full stack data practitioner through multiple software eras
- Been building data tools for 2 decades
- Built extensive/advanced RAG-type features
- Early adopter of MCP and its promises
- Learned a lot through building text-to-SQL @ Preset

Companion blog post @ preset.io

AI in BI: the Path to Full Self-Driving Analytics

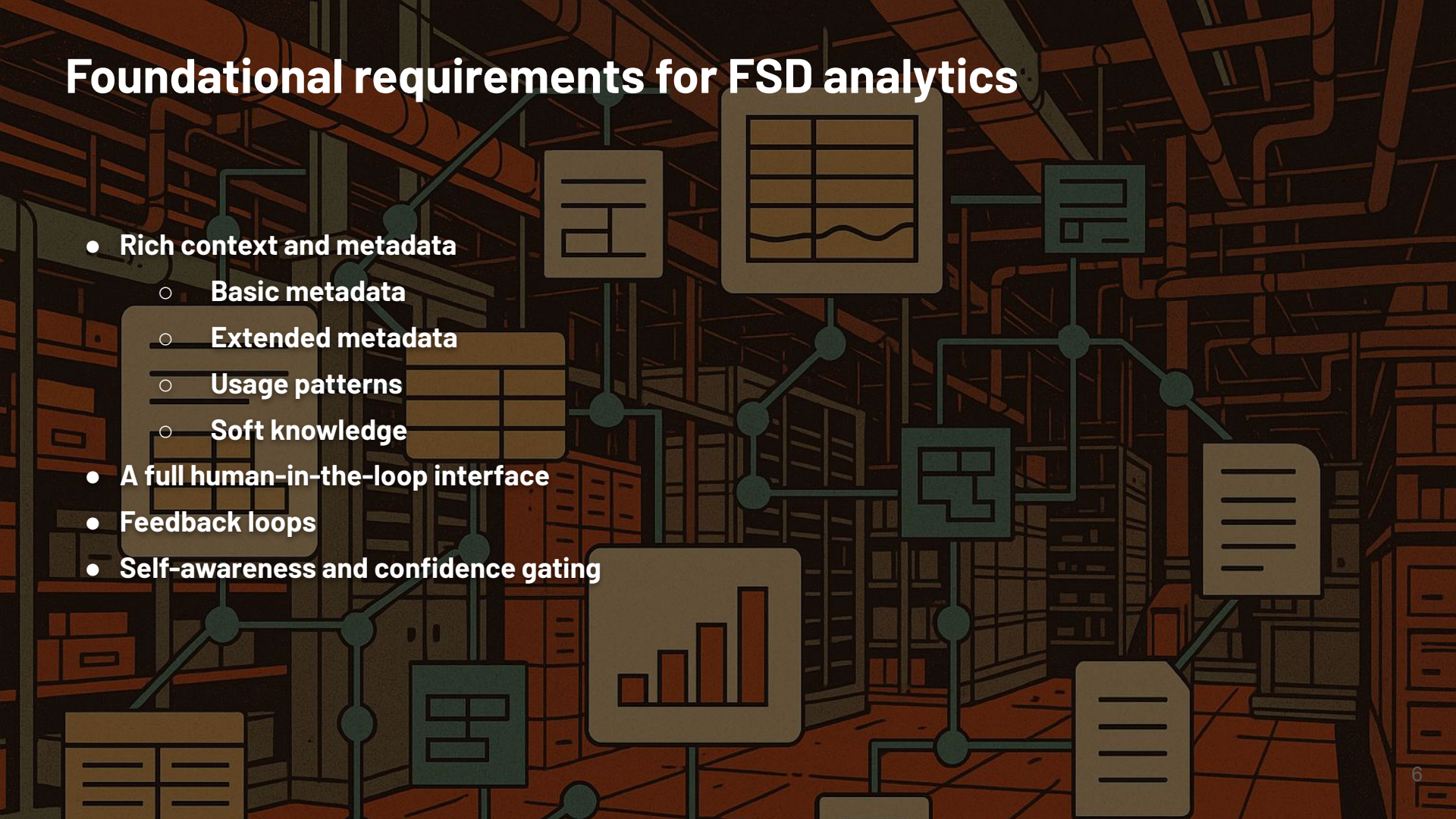
Maxime Beauchemin — SUPERSET DEVELOPERS

Two families of use cases



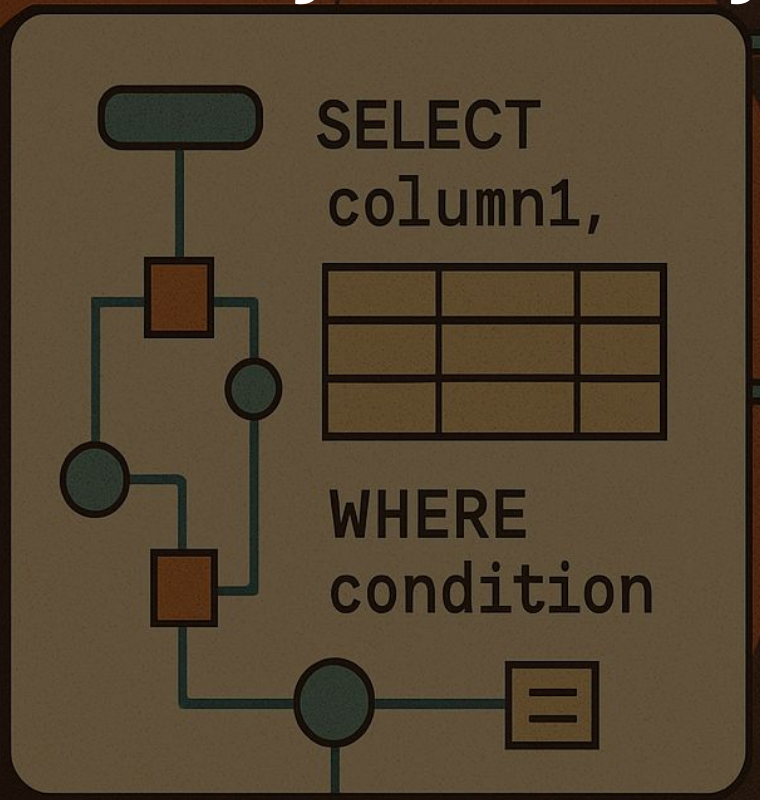
The lane assist vs. L5 analogy

Foundational requirements for FSD analytics

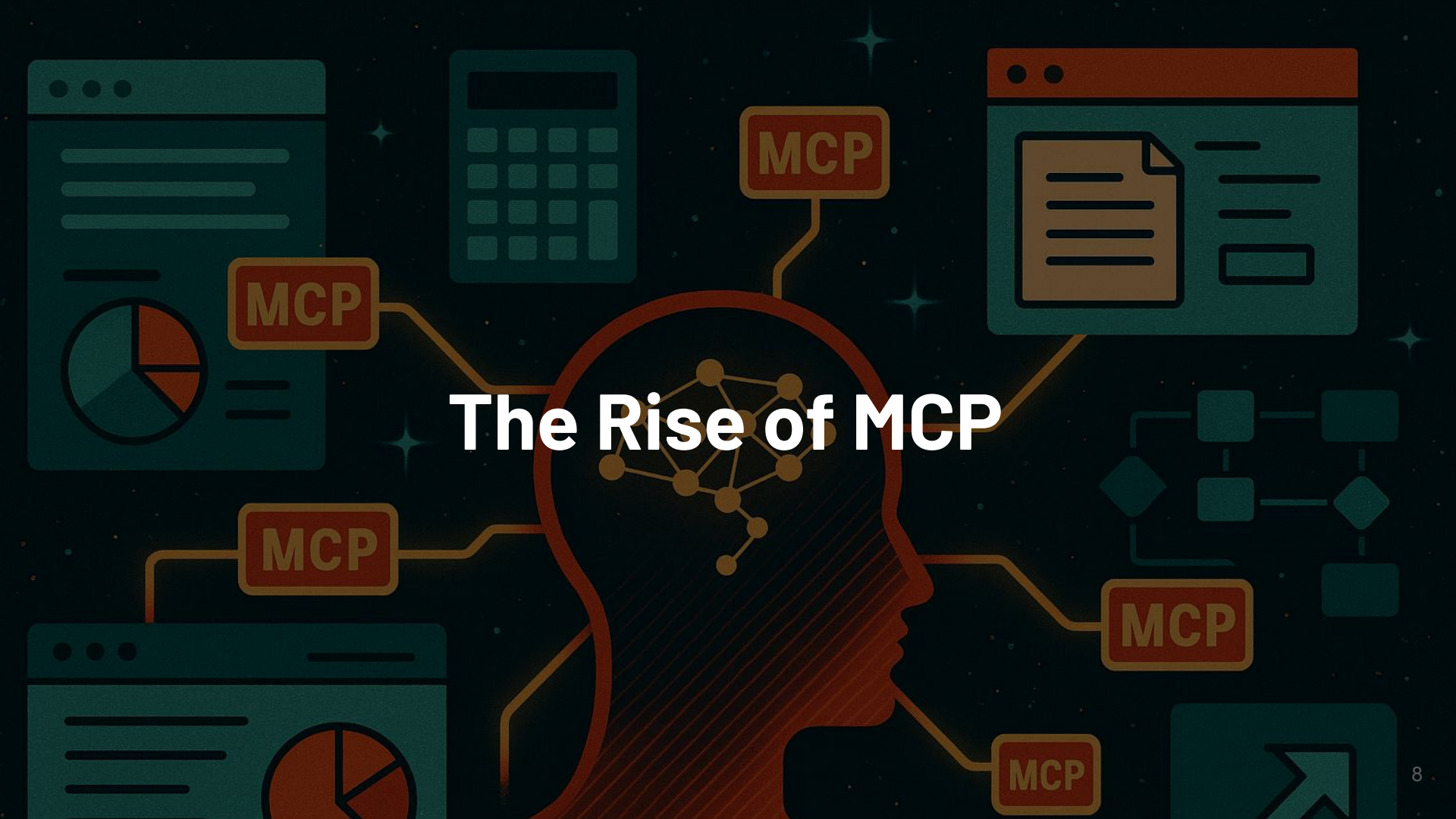


- Rich context and metadata
  - Basic metadata
  - Extended metadata
  - Usage patterns
  - Soft knowledge
- A full human-in-the-loop interface
- Feedback loops
- Self-awareness and confidence gating

Learnings while building text-to-SQL @ Preset



- Custom training & fine-tuning aren't viable RN
- RAG can be great — but has clear limitation
- Privacy and RBAC are real concerns in the way of adoption
- Magic moments exist — but trust is fragile
- Real-world data warehouses are messy AF
- Until 99% accuracy, introspection and a professional are REQUIRED



The Rise of MCP

THE SEMANTIC LAYER AS GUARDRAILS

Table

Tab

SEMANTIC LAYER

AI

EMPICIT

PUBLIC

KEY TAKEAWAYS

- The revolution will happen on familiar grounds - steering wheels will remain in car for quite some time
- AI entrustment will be a journey with a long tail - practitioners are likely to be behind the wheel to oversee AI
- Lookout for MCP! RAG to concede the place except for targeted, deep-in-product assisted workflows