RAG & SEARCH

ODSC AI

THE #1 AI TRAINING CONFERENCE

SAN FRANCISCO | OCT 28-30

BUILDING DYNAMIC IN-CONTEXT LEARNING FOR SELF-OPTIMIZING AGENTS

Amy Hodler

GraphGeeks

David Hughes

ENTERPRISE KNOWLEDGE

The State of Agent Observability Today

Blind optimization

Wasted resources

Arbitrary decisions

User frustration

We Demand More of Production Systems

Reliability: Detect failures before users do

Performance: Optimize latency & cost across workflows

Quality: Measure accuracy improvements

Compliance: Audit trail for regulated industries

Learning: Data for continuous improvement

There are options but we need more...

arize

opik by comet

mlflow™

A Modest Proposal

Observe

- **Comprehensive:** Cover all workflow components
- **Real-time:** Immediate feedback for production issues
- **Contextual:** Understand the "why" behind metrics
- **Actionable:** Clear path from insight to improvement

Guardrails

- **Aligned**: Business use cases, policies
- **Safety**: Ensure moderation
- **Actions**: What happens when a policy is triggered
- **Regulatory**: Follow domain requirements
- **Reporting**: Track violations for continuous learning

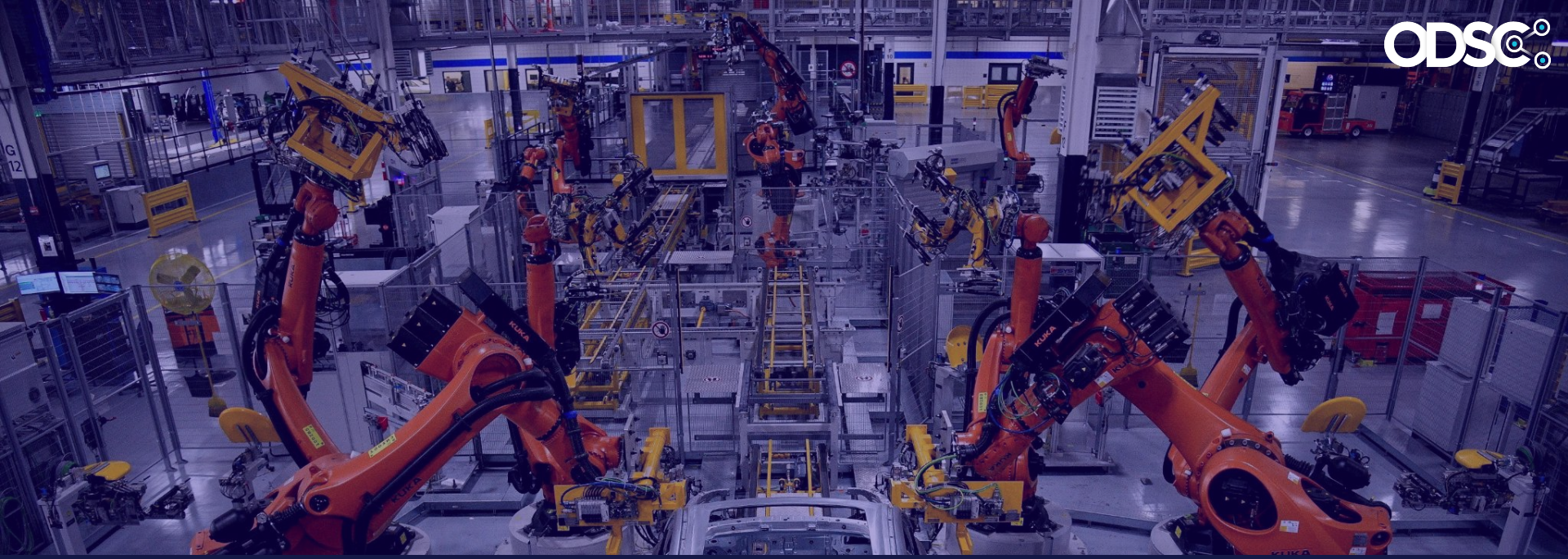
Evaluate

- **System Level:** Latency, throughput, resource usage
- **Workflow Level:** Success rates per agent/tool
- **Query Level:** Individual request tracing
- **Component Level:** Schema extraction, cypher generation quality

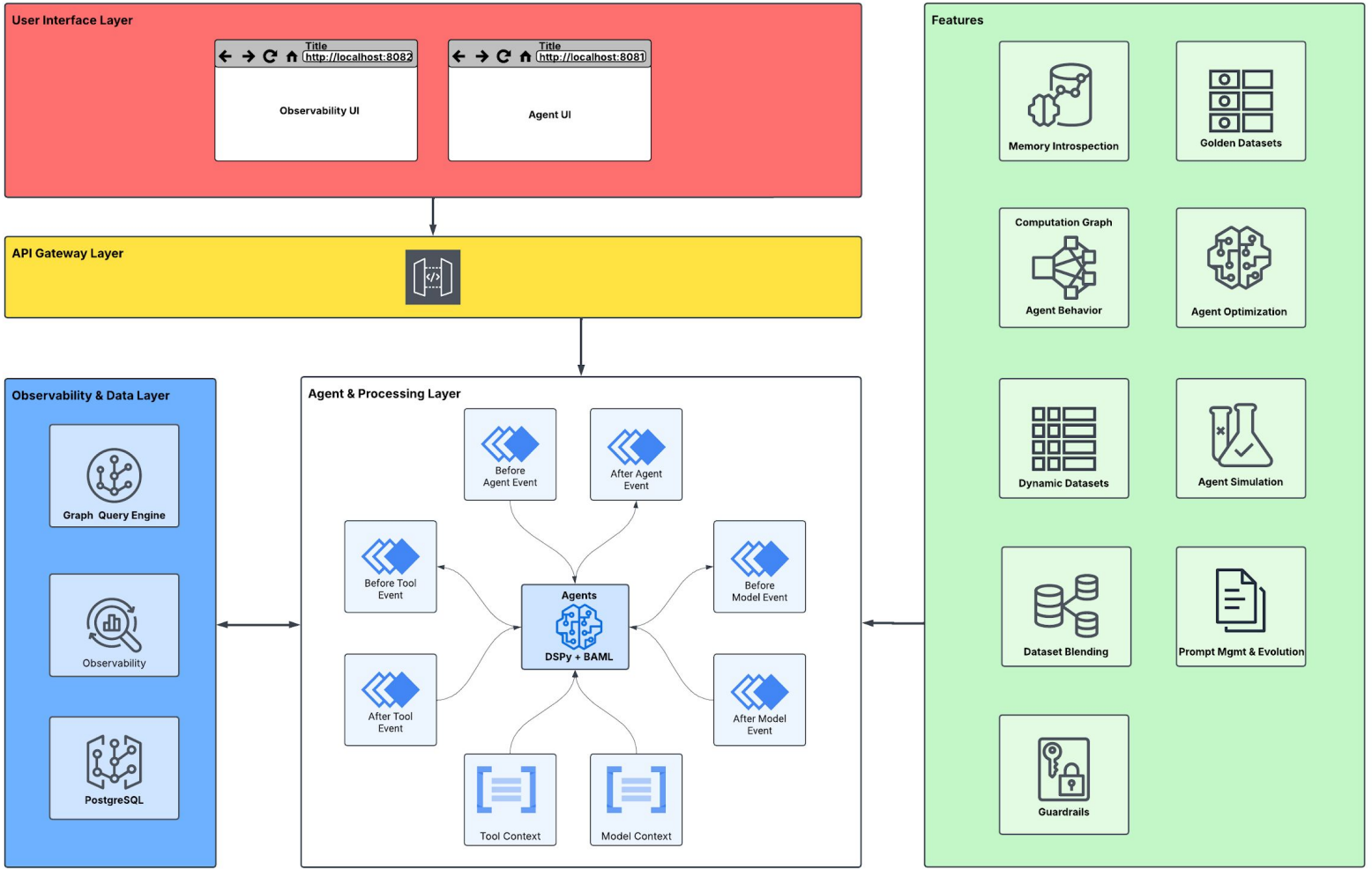
Optimize

- **Data**: Categorized training datasets
- **Heuristics**: Strategies for optimizations
- **Learn**: Successful strategies
- **Iterate**
- **Enable**: HITL & Self-Optimization

DEMO



Architecture


User Interface Layer
Title
http://localhost:8082
Observability UI
Title
http://localhost:8081
Agent UI
API Gateway Layer
Observability & Data Layer
Graph Query Engine
Observability
PostgreSQL
Agent & Processing Layer
Before Agent Event
After Agent Event
Before Tool Event
Agents
DSPy + BAML
Before Model Event
After Tool Event
After Model Event
Tool Context
Model Context
Features
Memory Introspection
Golden Datasets
Computation Graph
Agent Behavior
Agent Optimization
Dynamic Datasets
Agent Simulation
Dataset Blending
Prompt Mgmt & Evolution
Guardrails

Future Tenants

- Inspectability: Agent internals
- Behavior: What choices?
- Computation: Sequences
- Reasoning: What’s the rationale?
- Memory: How is past knowledge utilized & reinforced
- Simulation: How will perturbations affect AI solutions

Towards Self-Optimizing Agents

- Use observability and evaluation data for **automatic improvements**
- Feedback loops for **continuous learning**
- A/B testing, **Simulation**, **Experiments**, for workflow optimization
- **Adaptive routing** based on performance data
- What can be optimized?

Key Takeaways

- **Observability** & **Evaluation** is essential for production agents
- Data-driven optimization beats intuition
- Start with automatic instrumentation, add OOTB metrics

Our Team *is developing Self-Optimization Strategies for Agents using the capabilities of* ***Agents ↔ Observability****. Follow us to see our progress, learn, see us at upcoming events and presentations.*

Reach Out!

GraphGeeks.org

Join our vendor neutral community for graph practitioners & enthusiasts

Amy Hodler

Please Reach Out

Q&A

David Hughes

ENTERPRISE KNOWLEDGE