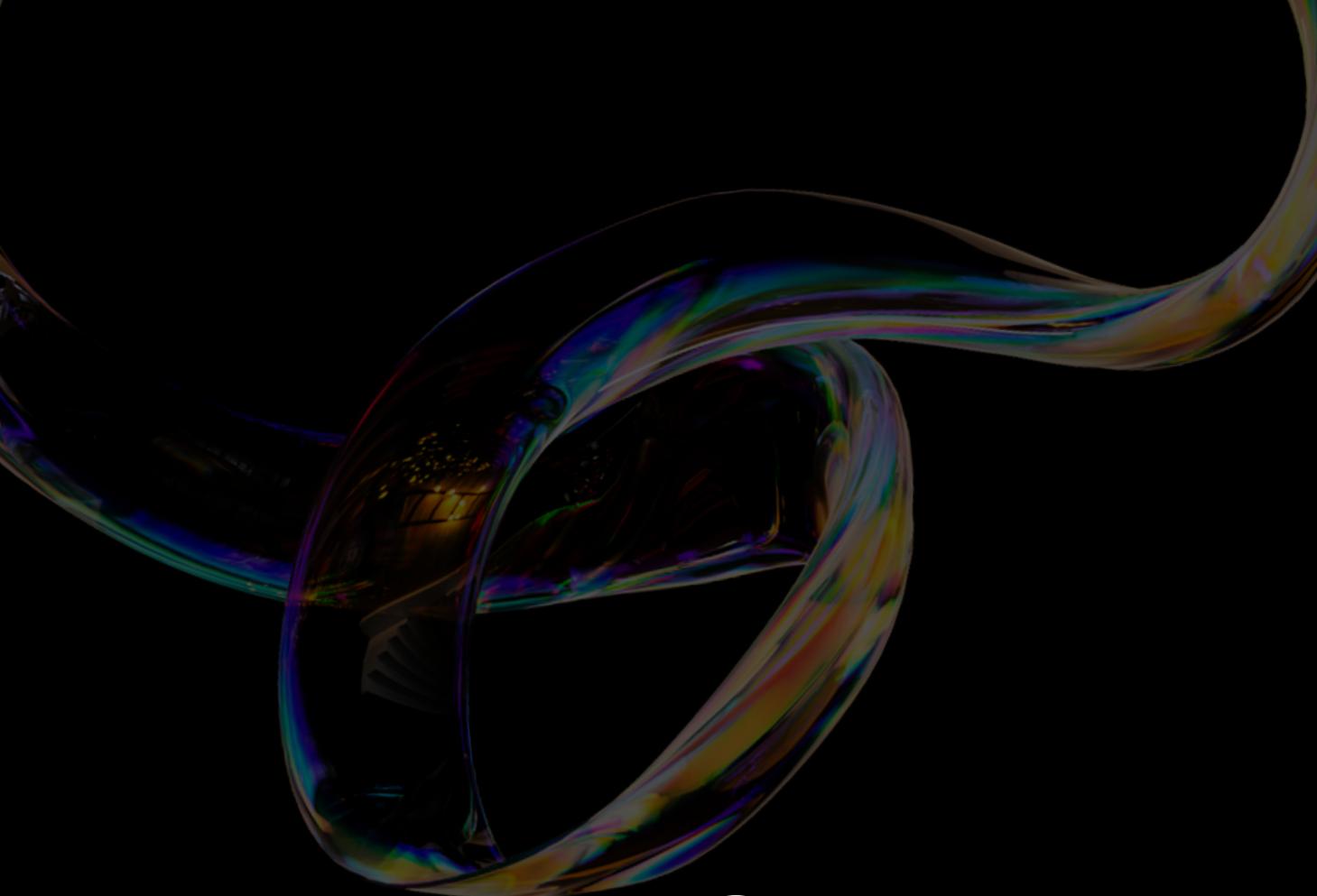


ADAPTIVE SELF- LEARNING AGENTIC AI SYSTEM

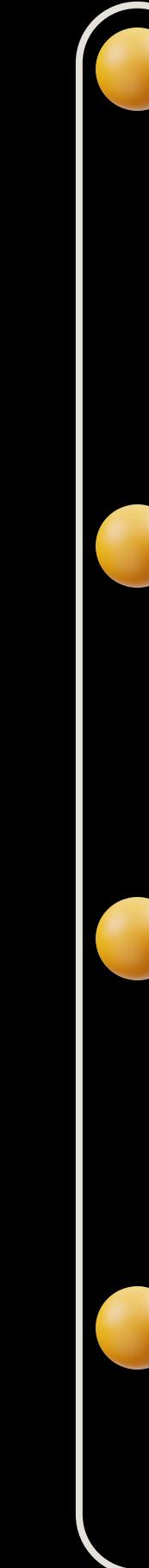
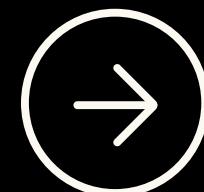
Continuous Fine-Tuning Framework
for Speech-to-Text Models



Kavya Venkatesh • Gautam Agarwal • Shivangi Kumar
Columbia University | COMS 6998 LLM-based Generative AI



THE PROBLEM: STATIC MODELS STAGNATE



Modern AI models deployed as static systems; they don't learn from mistakes

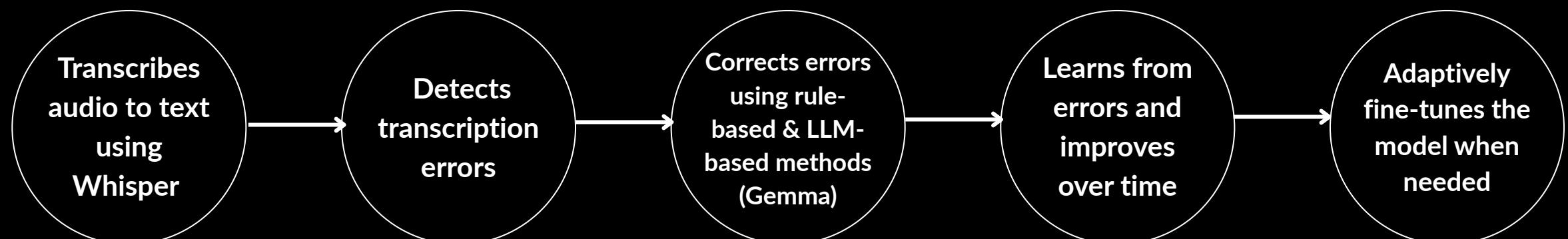
Performance degrades over time as domains shift

Manual retraining is expensive, slow, and reactive

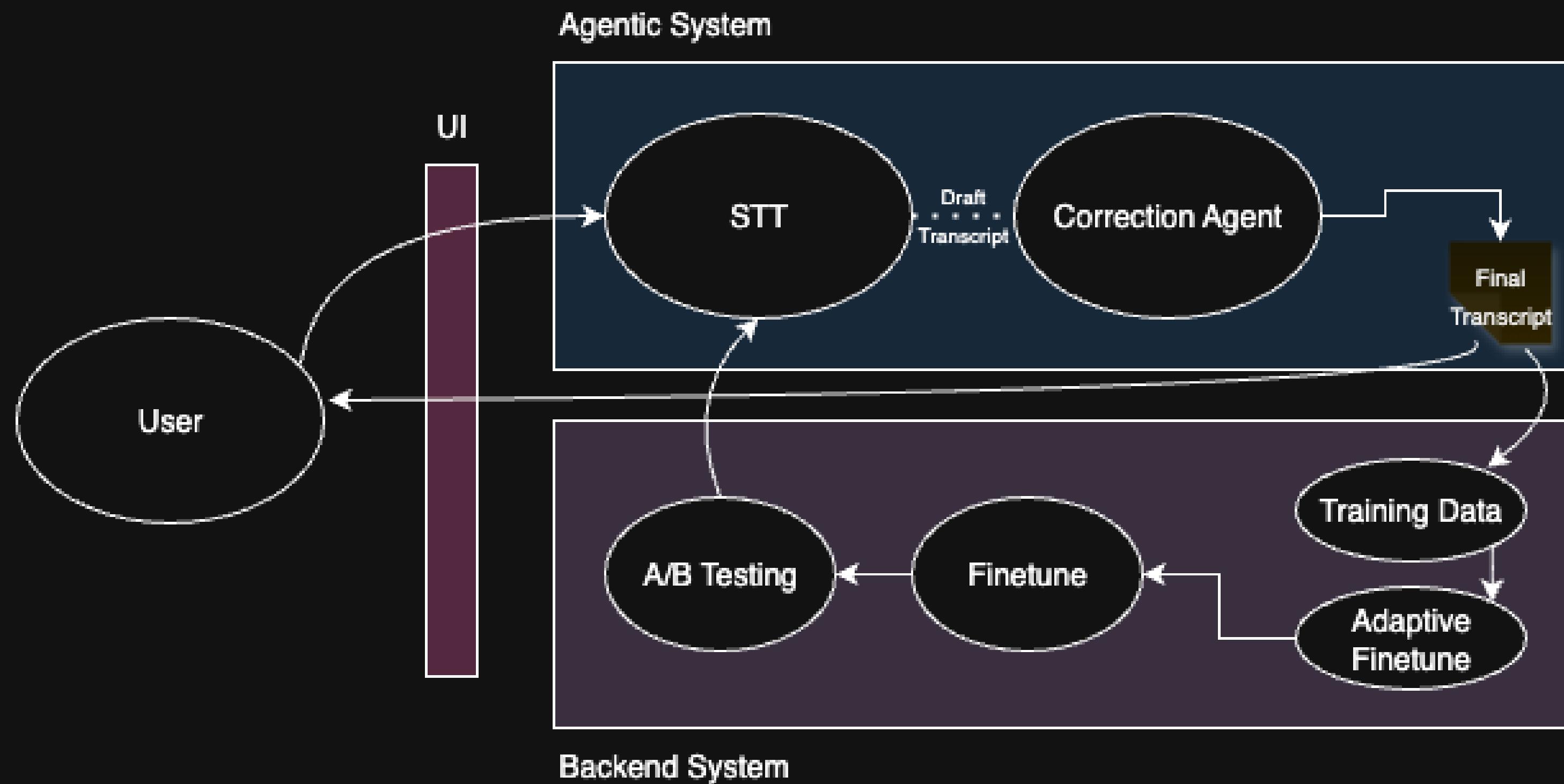
No closed-loop feedback: corrections happen, but models never improve

OUR SOLUTION: AUTONOMOUS SELF-IMPROVEMENT

- **Closed-loop system** detect errors → correct → retrain → improve
- **LLM-based correction agent** autonomously flags and fixes mistakes
- **Automated fine-tuning pipeline** with hyperparameter optimization
- **Adaptive scheduling** prevents overfitting and controls cost



ARCHITECTURE DIAGRAM



EVALUATION: TASK-BASED METRICS

Word Error Rate

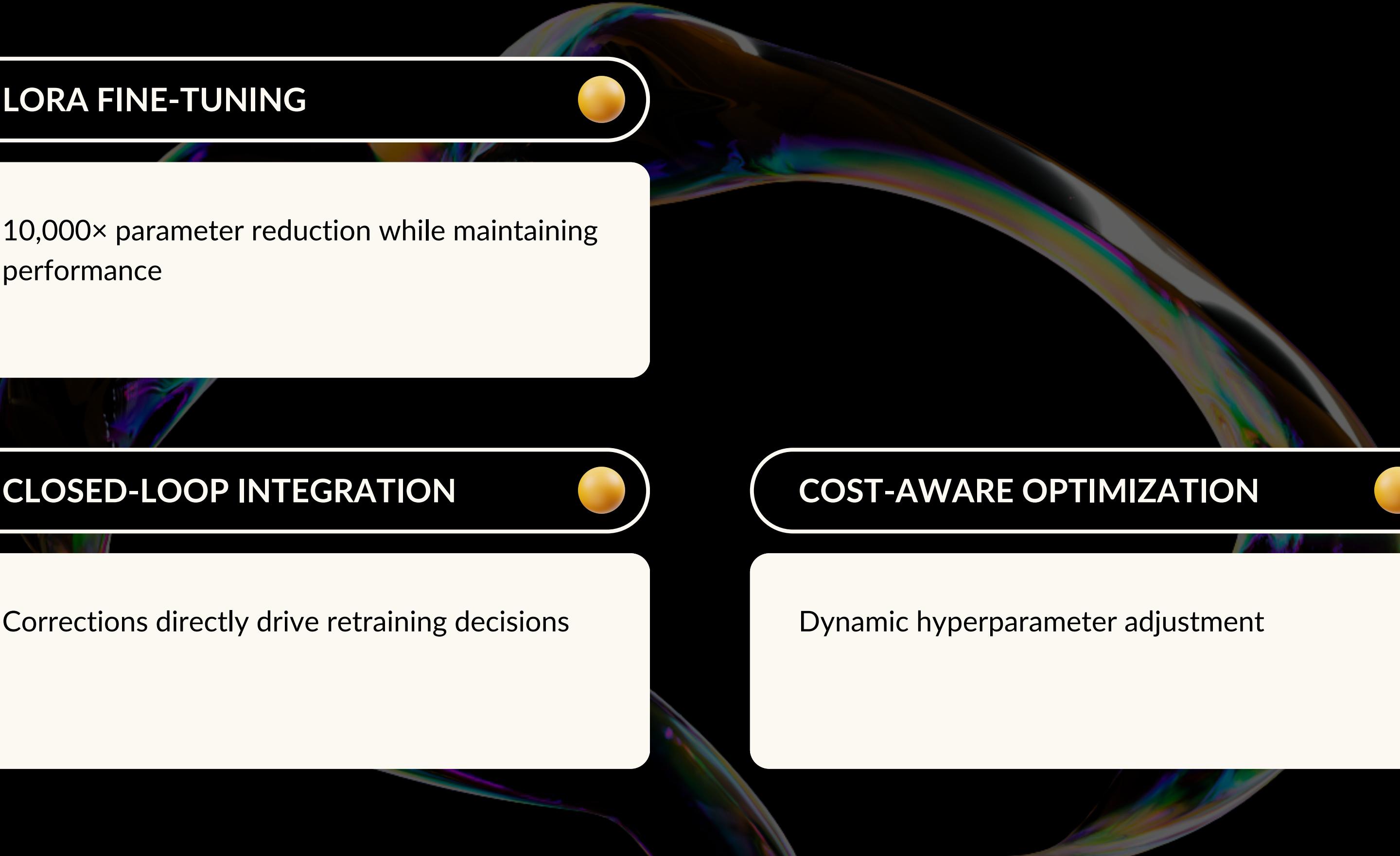
Character Error Rate

Error Detection Precision/Recall

Latency, Throughput, GPU Utilization

Training & Inference Cost

KEY TECHNICAL INNOVATIONS



LORA FINE-TUNING

10,000× parameter reduction while maintaining performance

CLOSED-LOOP INTEGRATION

Corrections directly drive retraining decisions

COST-AWARE OPTIMIZATION

Dynamic hyperparameter adjustment

EXPECTED OUTCOMES



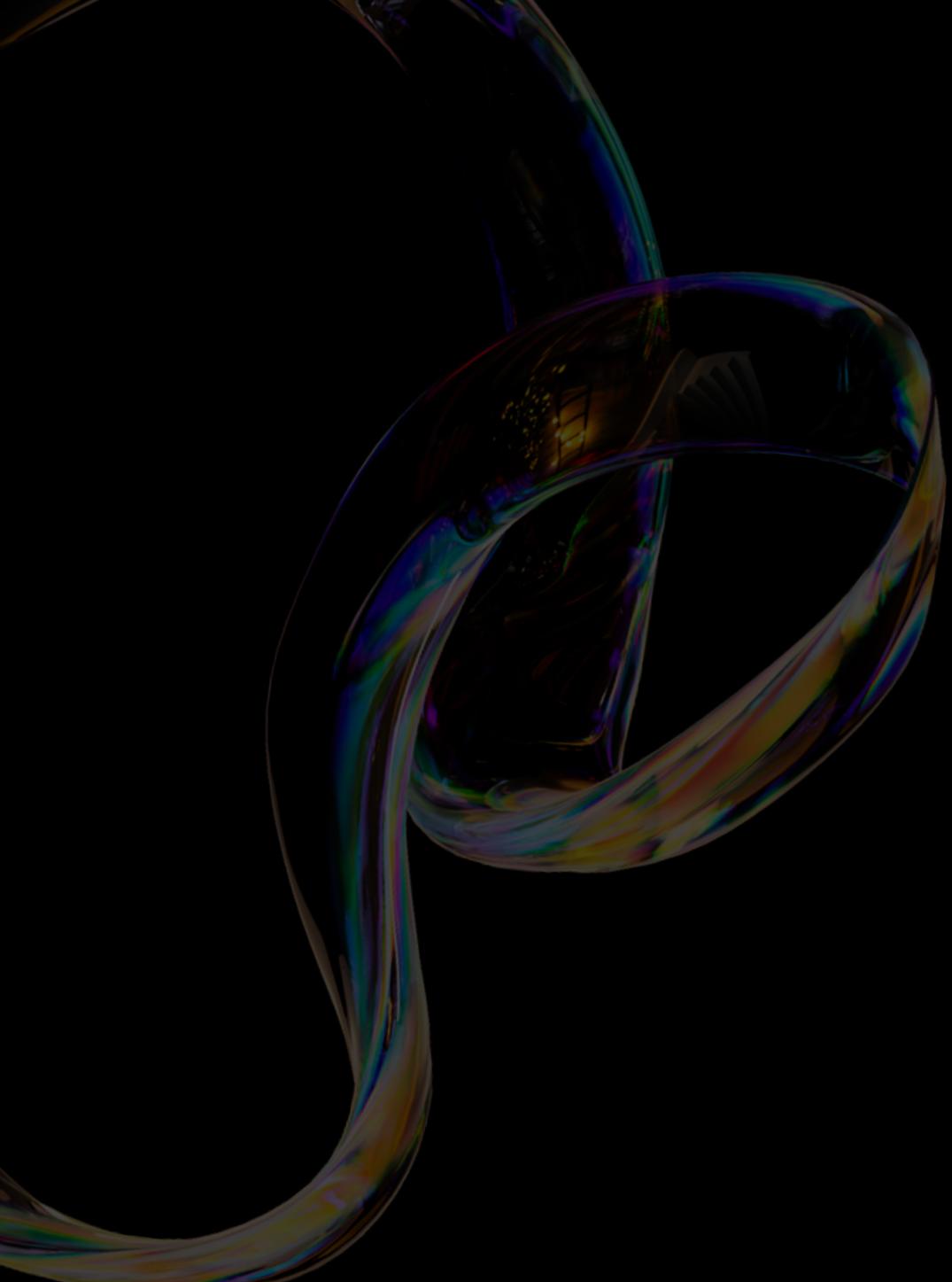
Functional self-improving STT system with measurable accuracy gains

Catastrophic forgetting prevention using replay-based methods & parameter regularization

Fully automated hyperparameter tuning and retraining pipeline

Adaptive scheduling that balances performance and computational cost

Generalized framework applicable to multiple generative AI domains



TIME FOR A DEMO

RISK MITIGATION STRATEGY

Overfitting

Adaptive scheduling + validation sets

Resource Cost

LoRA lightweight fine-tuning

Degradation

Regression testing before deploy

IMPACT + SIGNIFICANCE

First unified closed-loop system combining correction and adaptive fine-tuning

Reduces manual intervention and retraining costs by orders of magnitude

Enables zero-touch continuous improvement for production ML systems

Generalizable framework applicable across all generative AI domains

FUTURE DIRECTION

We're here!



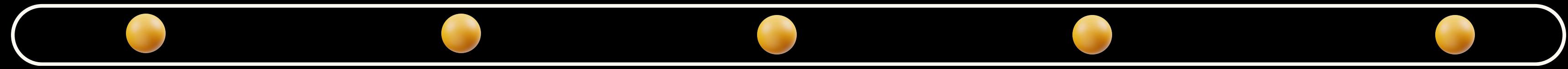
Visualization
dashboards for
monitoring

Performance
benchmarking on
larger datasets

Better Adaptive
Decision-making
Strategy

Framework
Generalization

Human-in-the-loop
IOption Integration



THANK YOU

<https://github.com/GAIInTheHouse/Adaptive-Self-Learning-Agentic-AI-System>



Kavya Venkatesh • Gautam Agarwal • Shivangi Kumar
Columbia University | COMS 6998 LLM-based Generative AI