

Data-Centric AI

Alexander Ratner

Founder & CEO, Snorkel AI



Thank you for coming

JP Morgan Chase & Co. | Barclays | Pfizer | Comcast | Experian |
State Street | NASA | Wayfair | Jeffries | The Walt Disney Company

Data-centric AI starts here

Unlock the Future of Enterprise AI

 **Oct 16-17, 2024** | New York, NY

**Sold
Out!**

Thank You to Our Sponsors



Google Cloud



Microsoft



databricks



SambaNova[®]
SYSTEMS



Hands-On Training Day

Join our instructor-led workshop where you will get hands-on experience using Snorkel Flow to build AI applications rapidly using a data-centric AI workflow, weak supervision, and programmatic labeling.

Who should attend?

Data scientists, machine learning, and software engineers within Global 2000 enterprises and government organizations.



Introducing **Snorkel** **Flow R3**

One enterprise platform for data development across predictive and generative models

- New GenAI features
- New NLP features
- Efficient annotation

SnorkelCon²⁴ New York City

8:00am

Registration & Breakfast

9:00am

Opening Keynote: Bridging the Gap Between LLMs and Enterprise AI

Alex Ratner, Co-founder and CEO, Snorkel AI

9:45am

Snorkel Research Spotlight

Fred Sala, Professor of CS, University of Wisconsin & Chris Glaze, Principal Research Scientist, Snorkel AI

10:15am

Coffee Break

10:30am

Tech Session: Graduating from Data Labeling to AI Data Development

Chris Glaze, Principal Research Scientist, Snorkel AI & Angela Fox, Staff Product Designer, Snorkel AI

Business Session:

Experian: The Importance of LLM Evaluation for Domain-Specific Use Cases

James Lin, Head of AI/ML Innovation, Experian

11:20am

Tech Session:

Unlocking Hidden Insights: Snorkel's Solutions for Complex and High-Value PDF Documents

Jennifer Lei, Senior Product Manager, Snorkel AI

Business Session:

How Citi is Succeeding with AI in Banking

Aarti Bagul, Head of Field Engineering, Snorkel AI & Femi Agboola, Managing Director, Citi Productivity, Citi

12:05pm

Lunch brought to you by Microsoft

1:00pm

Tech Session:

Evaluating LLM Systems

Rebekah Westerlind, Software Engineer, Snorkel AI & Venkatesh Rao, Staff Product Manager, Snorkel AI

Business Session:

How Wayfair is Transforming Customer Experience with Data-Centric AI

Vinny DeGenova, Associate Director of Machine Learning, Wayfair

1:50pm

Tech Session:

Enhancing RAG Pipelines for Enterprise-Specific Tasks: Strategies for Accuracy and Reliability

Bryan Wood, Machine Learning Solutions Engineer, Snorkel AI

Business Session: AI From the Trenches: Lessons Learned from Practitioners on the Front Lines

Alex Shang, Machine Learning, Snorkel AI, André Balleyguier, Head of ML Field Engineering, Snorkel AI, Elena Boiarskaia, Head of Applied Machine Learning, Snorkel AI, Gabe Smith, Senior Machine Learning Success Manager, Snorkel AI

2:40pm

Tech Session:

Fine-Tuning and Aligning LLMs with Enterprise Data

Marty Moesta, Lead Product Manager, Generative AI, Snorkel AI & Amit Kushwaha, Director of AI Engineering, SambaNova Systems

Business Session:

Delivering Business Value with Data-Centric AI in Financial Services

Peter Williams, Head of Partner Technology, Global Financial Services, AWS & Bryan Wood, Machine Learning Solutions Engineer, Snorkel AI

3:45pm

Product Keynote:

The Future of Snorkel AI

Ronaldo Ama, Chief Development Officer, Snorkel AI & Ajay Singh, Chief Product Officer, Snorkel AI

4:30pm

Closing Keynote:

Snorkel and Citi discuss the future of AI in the enterprise

Alex Ratner, Co-founder and CEO, Snorkel AI & Murli Buluswar, Head of Analytics, US Personal Bank, Citi

5:15pm

Mix and Mingle

Cocktails and hors d'oeuvres

5:45 pm

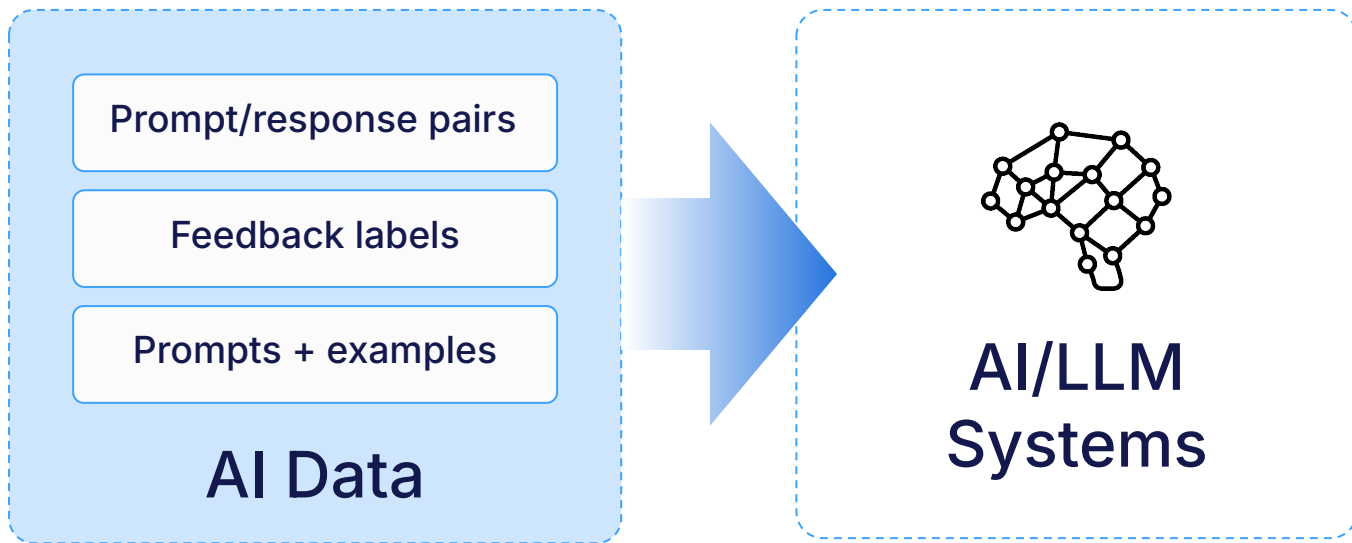
Magic Show featuring World Renowned Magician Alexander Boyce

Cocktails and hors d'oeuvres will continue

6:15 pm

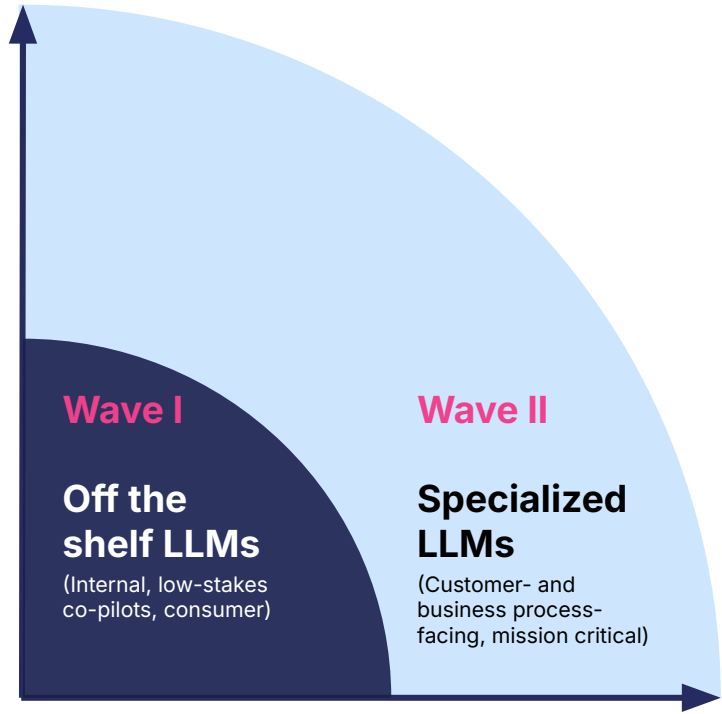
Event Appreciation Party

Current State: Data as the Interface for AI



Data is the interface for AI

Unique domain & organizational specific requirements



Wave I

Wave II

Off the shelf LLMs

(Internal, low-stakes co-pilots, consumer)

Specialized LLMs

(Customer- and business process-facing, mission critical)

Accuracy requirements

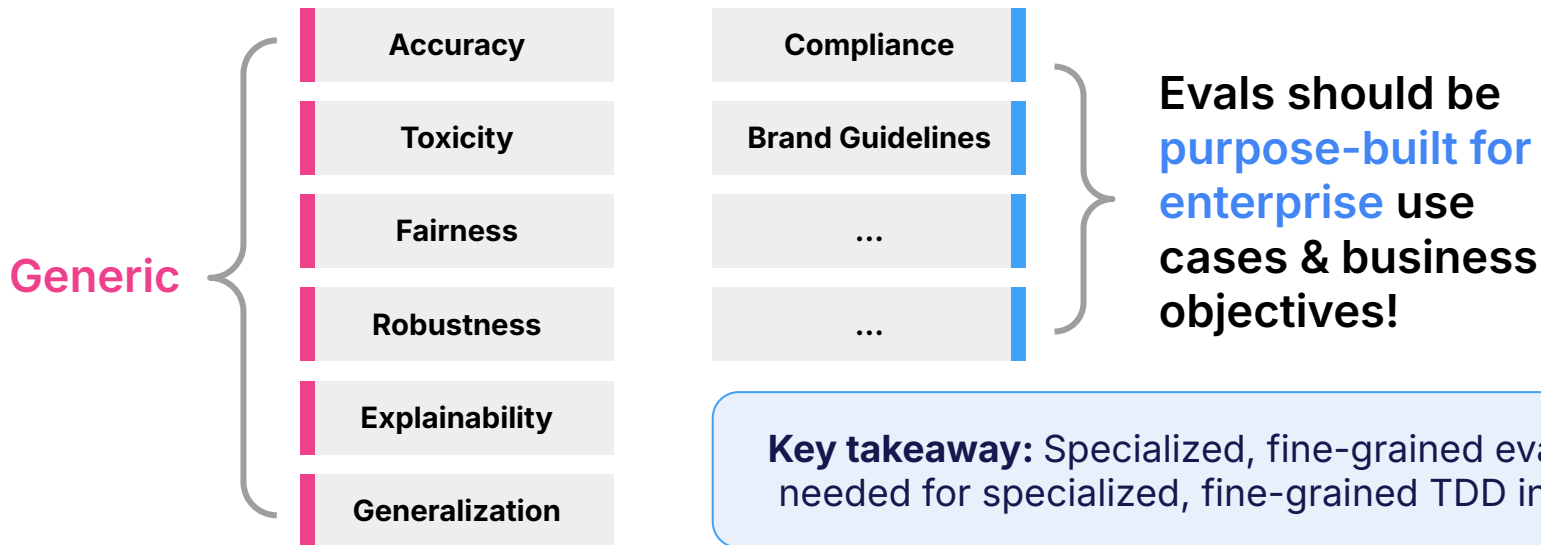
Specialization required for higher ROI "Wave II"

Critical Themes Ahead

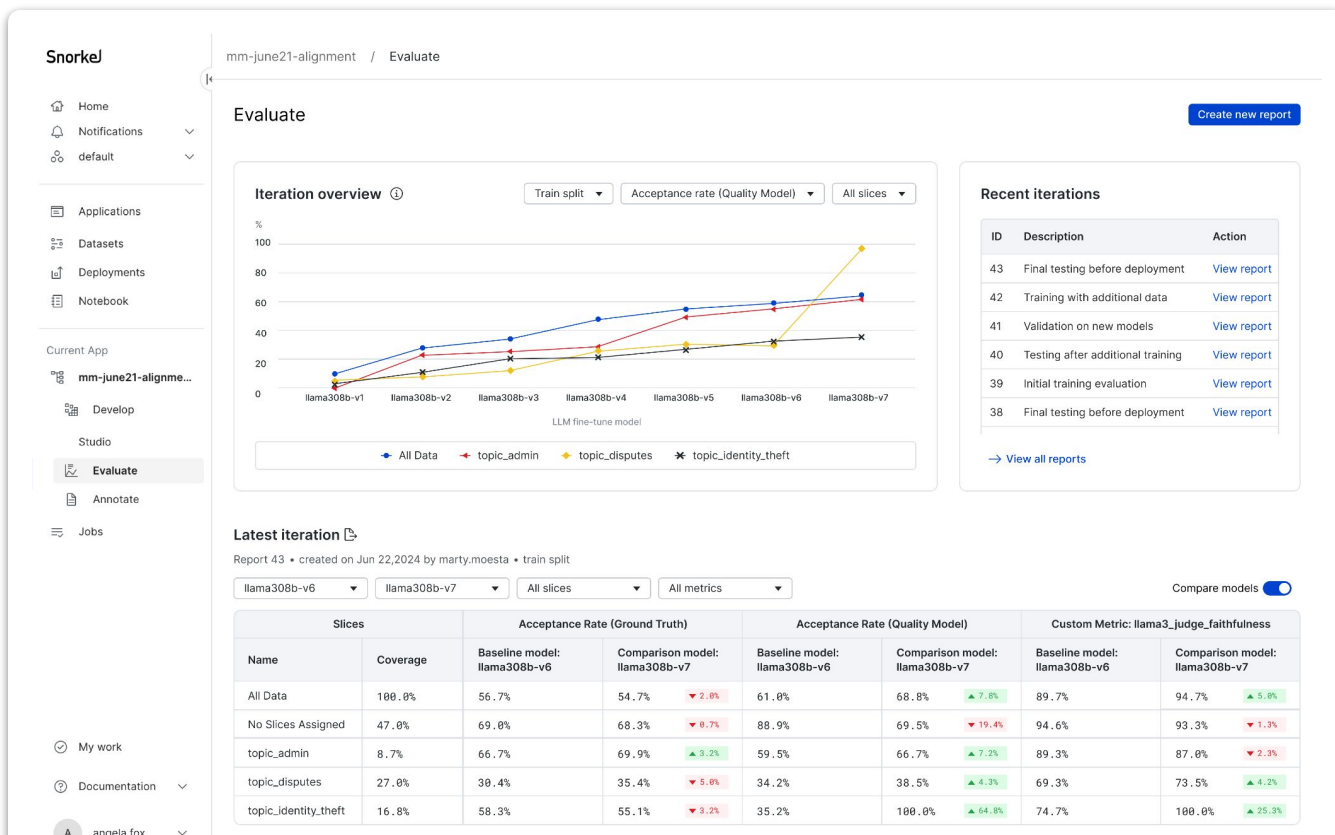
- Specialized, fine-grained, scalable evaluation
- Full LLM system tuning/specialization
- A convergence of data-centric interfaces
- Full lifecycle human feedback as a key differentiator



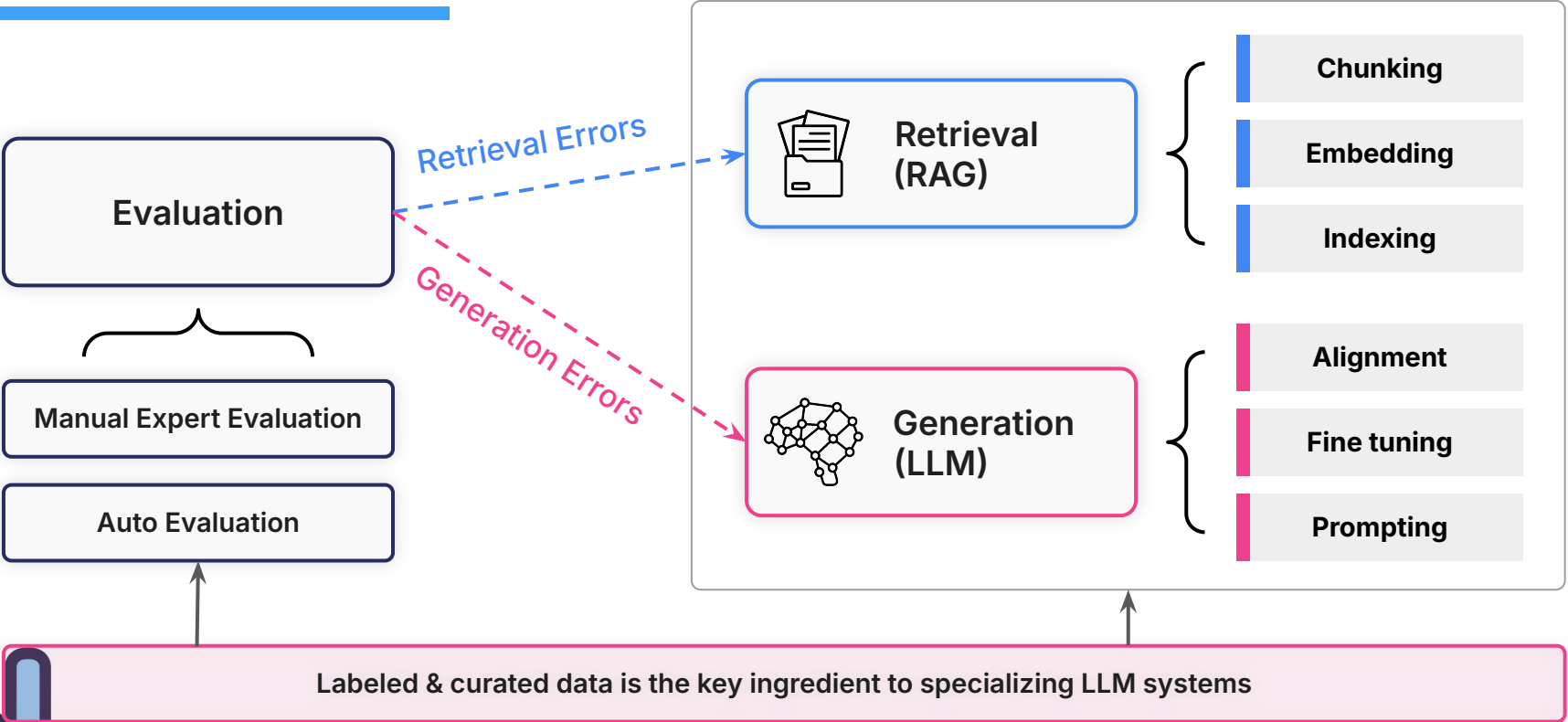
Specialized, fine-grained, scalable evaluation as the starting point and epicenter



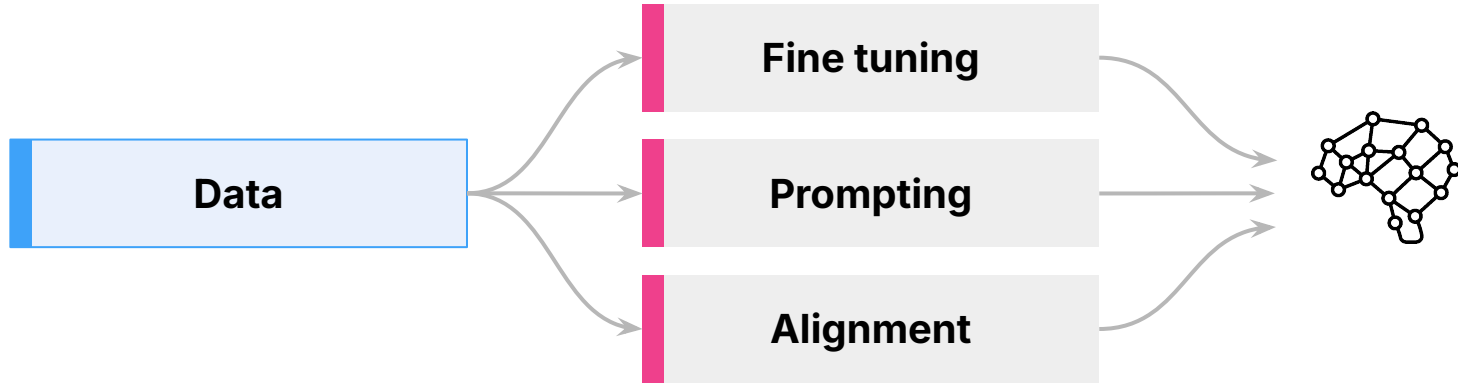
Fine-grained, specialized evaluations for test-driven AI development



From evaluation to LLM system tuning



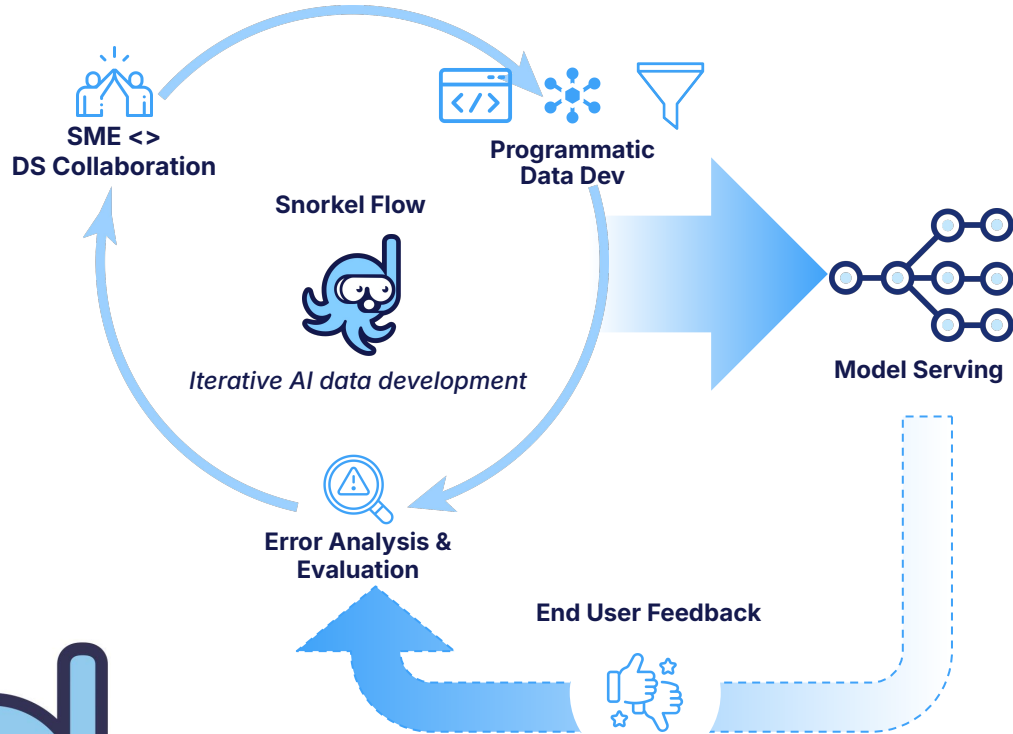
Convergence of data-centric interfaces



As methods for specializing AI converge,
data is the key ingredient to focus on



Capturing and using full-cycle human feedback



Key takeaway: Labeled data for tuning/evaluating/specializing comes throughout the deployment lifecycle

Looking Ahead: The Next Steps for Data-Centric AI



Data as the common *transportable* artifact across models & interfaces

- 1 Data for tuning and aligning multi-component systems → agentic systems
- 2 Data for tuning inference-time reasoning (e.g. o1-style)

Our aspiration, together with all of you:

**Make AI data development
programmatic to power the next wave
of enterprise AI innovation**



SnorkelCon²⁴
New York City