Speakers:

Sandeep Jha, Principal Staff Technical Program Manager, **LinkedIn**
Aman Gupta, Sr. Staff Engineer - Applied AI Research, **LinkedIn**
Tejas Chopra, Senior Software Engineer, **Netflix**
Saurabh Vishwas Joshi, Tech Lead, Senior Staff Engineer - ML Platform, **Pinterest**
Elsa Jungman, CEO and Founder, **HelloBiome**
Prithvishankar Srinivasan, Machine Learning Lead, **Instacart**
Shih-Ting Lin, Senior Machine Learning Engineer, **Instacart**
Sai Kumar Arava, Machine Learning Manager, Gen AI / ML Applications, **Adobe**
Sayan Maity, Principal ML Engineer, **Disney Streaming**
Faith McGee, Sr. AI/Analytics Research Manager, **Electronic Arts**
Natashia Tjandra, Research Director, **Electronic Arts**
Jitender Jain, Lead Software Engineer, **ex-Walmart/ Capital One**
Preetish Kakkar, Senior Computer Graphics Engineer, **Adobe**
Abhai Pratap Singh, Senior Product Manager-Technical, **Amazon Alexa**
Ipsita Basu, Product Lead, **Shopify**
Lisa M. Lum, Founder & CEO, **Friends of the Metaverse**
Chaitanya (Chai) Pydimukkala, Product Head, BigQuery Governance, **GoogleCloud**
Muhammad Farooq, Manager of Machine Learning, **MindsDB**

Day 1 – 11 February **Focus:** Advances in Deep Learning and Real-World AI Applications	
8:15-9:00	**REGISTRATION AND NETWORKING BREAKFAST**
9:00-9:10	**CHAIR OPENING REMARKS** Lisa M. Lum, Founder & CEO, Friends of the Metaverse
9:10-9:50	**Opening Panel: Embedding Ethics in AI Development – Best Practices for Fair and Accountable Systems** • Discuss the ethical considerations surrounding AI development and deployment. • Learn about best practices for building fair, transparent, and accountable AI systems. • Explore the role of governance and regulation in promoting ethical AI. Mod: Lisa M. Lum, Founder & CEO, Friends of the Metaverse 1 – Jitender Jain, Lead Software Engineer, JPMorgan Chase 3 – Sayan Maity, Principal ML Engineer, Disney Streaming 4 - Abhai Pratap Singh, Senior Product Manager-Technical, Amazon Alexa
9:50-10:20	**Keynote Session Presented by GoogleCloud** To follow … Chaitanya (Chai) Pydimukkala, Product Head, BigQuery Governance, GoogleCloud
10:20-10:50	**Case Study: AI-Driven Microbiome Insights for Personalized Care and Longevity** To follow … Elsa Jungman, Founder & CEO, HelloBiome
10:50-11:10	**COFFEE AND NETWORKING BREAK**
11:10-11:40	**Lessons from Building Enterprise RAG System** • Key insights and strategies from working with enterprise clients to build effective retrieval-augmented generation (RAG) systems. • Practical tips for designing “Minds” for optimal retrieval performance. • What worked, what didn’t, and how to tackle common challenges. Muhammad Farooq, Manager of Machine Learning, MindsDB

11:40-12:10	**Memory Optimizations in Machine Learning** As Machine Learning continues to forge its way into diverse industries and applications, optimizing computational resources, particularly memory, has become a critical aspect of effective model deployment. This session, "Memory Optimizations for Machine Learning," aims to offer an exhaustive look into the specific memory requirements in Machine Learning tasks, including Large Language Models (LLMs), and the cutting-edge strategies to minimize memory consumption efficiently. We'll begin by demystifying the memory footprint of typical Machine Learning data structures and algorithms, elucidating the nuances of memory allocation and deallocation during model training phases. The talk will then focus on memory-saving techniques such as data quantization, model pruning, and efficient mini-batch selection. These techniques offer the advantage of conserving memory resources without significant degradation in model performance. A special emphasis will be placed on the memory footprint of LLMs during inferencing. LLMs, known for their immense size and complexity, pose unique challenges in terms of memory consumption during deployment. We will explore the factors contributing to the memory footprint of LLMs, such as model architecture, input sequence length, and vocabulary size. Additionally, we will discuss practical strategies to optimize memory usage during LLM inferencing, including techniques like model distillation, dynamic memory allocation, and efficient caching mechanisms. By the end of this session, attendees will have a comprehensive understanding of memory optimization techniques for Machine Learning, with a particular focus on the challenges and solutions related to LLM inferencing. Tejas Chopra, Senior Software Engineer, Netflix
12:10-12:30	**AI as a Strategic Asset: Building Competitive Advantage in Financial Services and Enterprises** • Explore how financial services and enterprises can leverage AI as a strategic asset to gain a competitive edge. • Learn about tailored strategies for designing and implementing successful AI initiatives in financial services and broader business contexts. • Discover real-world examples of companies leveraging AI for strategic advantage. Jitender Jain, Lead Software Engineer, ex-Walmart/ Capital One
12:30-13:30	**NETWORKING LUNCH**
13:30-14:00	**Duo Presentation: Making Smaller LLMs Punch Above Their Weight - Lessons in Post-Training and Fine-Tuning**

	LLMs have started providing great value to enterprise use cases. Deploying smaller LLMs in production is attractive since they are less resource hungry. Smaller LLMs, though efficient, often struggle to match the performance of their bigger counterparts. In this talk, we will discuss various techniques like knowledge distillation, post-training alignment like RLHF and DPO, etc. that can help bridge this gap. The speakers will discuss success stories from open source and enterprise applications. Sandeep Jha, Principal Staff Technical Program Manager, LinkedIn Aman Gupta, Sr. Staff Engineer - Applied AI Research, LinkedIn
14:00-14:30	**Launching Gen AI Agents in Production: Best Practices for Enterprise Use Cases** • Understand the best practices for scaling, deploying, and managing Gen AI agents in enterprise environments. • Learn how to address challenges related to security, compliance, and performance optimization in production. • Gain actionable strategies for leveraging Gen AI to drive innovation, improve operational efficiency, and create business value. Sai Kumar Arava, Machine Learning Manager, Gen AI / ML Applications, Adobe
14:30-14:50	**Scaling Deep Learning-Based Recommender Model Training** • Acquire insights on operationalizing, optimizing, and efficiently scaling deep learning model training. • Learn from case studies, on managing ML platforms with web-scale data. • Understand how modern ML computing frameworks like Ray and PyTorch can be utilized to create impactful ML products. • Dive into deep technical discussions regarding ML/AI and infrastructure Saurabh Vishwas Joshi, Tech Lead, Senior Staff Engineer - ML Platform, Pinterest
14:50-15:10	**AFTERNOON NETWORKING BREAK**
15:10-15:40	**Understanding the Synergy Between 3D Graphics and Machine Learning** • AI-Driven 3D Content Creation: Showcase how machine learning algorithms automate and enhance 3D modeling (Discuss about NeRF) • AI-Driven Occlusion for XR Applications: Highlight how AI models, such as those derived from ARKit and ARCore, enable accurate object and people occlusion, which is critical for creating immersive and realistic XR experiences. • Deep Learning for Order-Independent Transparency (OIT): Explore how deep learning techniques can enhance traditional transparency algorithms like k-buffer, improving performance and visual quality in real-time applications where transparency is essential.

	Preetish Kakkar, Senior Computer Graphics Engineer, Adobe
15:40-16:10	**Level Up: How Gaming is Winning with AI** • Achieving the Unimaginable: See how AI enables the creation of thousands of real college football players in College Football 25, enhancing player experiences at scale. • Breaking Barriers with GPTs: Gain actionable strategies to expand into new markets and engage diverse audiences, including neurodiverse customers, using advanced generative AI tools. • Navigating AI Risks: Understand the pitfalls of bias in large AI models and learn best practices to mitigate its impact for more equitable and ethical outcomes. Faith McGee, Sr. AI/Analytics Research Manager, Electronic Arts Natashia Tjandra, Research Director, Electronic Arts
16:10-16:40	**Duo Presentation: GenAI Extraction for E-commerce at Instacart** Prithvi Srinivasan and Shih-Ting Lin will explore how Instacart leverages Large Language Models (LLMs) to extract and manage information from raw product data within Instacart's extensive catalog, which includes millions of products from stores across the US and Canada. Their presentation will cover the process of extracting essential product attributes, the effective utilization of these attributes, and the tooling developed for backend engineers to streamline these operations. Prithvi Srinivasan, Machine Learning Lead, Instacart Shih-Ting Lin, Senior Machine Learning Engineer, Instacart
16:40-17:10	**Closing Keynote: AI-Powered Personalization** Personalization has become the cornerstone of innovation across industries, driven by AI's ability to understand, anticipate, and respond to individual user needs at scale. From financial services creating tailored investment plans to healthcare providing personalized treatment recommendations, AI enables hyper-relevant experiences that delight users and deliver measurable results. This talk explores how AI-powered personalization is transforming the way businesses deliver value and build stronger connections with users across diverse sectors. Ipsita Basu, Product Lead, Shopify
17:10	**CLOSING REMARKS**
17:00-18:00	**EVENING RECEPTION**