





WHAT RECENT DEVELOPMENT IN THE AI INDUSTRY
ARE YOU MOST EXCITED ABOUT?

Large Agentic Models: Massive language models exhibiting human-like capabilities, serving as foundation models for various AI tasks.

Multi-modal AI: Systems processing and integrating information from multiple modalities (text, images, audio, video) for comprehensive understanding and cross-domain applications.

Mixture of Experts: Al architecture combining multiple expert models, each specializing in specific tasks or domains, enabling efficient scaling and leveraging diverse expertise.

Model Distillation and Edge Serving: Transferring knowledge from large models to smaller, efficient models for deployment on resource-constrained edge devices while maintaining reasonable performance.

Infrastructure optimization (GPU/CPU), workload management, model routing and selecting the right model for the right task.



SPEAKER SPOTLIGHT

Galina Sagan *Prinicpal* **Hitachi Ventures**





WHAT DO YOU THINK IS GOING TO BE THE BIGGEST PROBLEM ENTERPRISES WILL FACE IN MANAGING THEIR AI INFRASTRUCTURE IN THE NEXT YEAR?

Scalability and Efficiency: Handling complex AI models and large (especially unstructured) datasets efficiently.

Infrastructure Costs: Managing expenses associated with hardware and cloud resources.

Data Management: Organizing, preprocessing, and ensuring the quality and security of large datasets.

Model Training and Deployment: Optimizing training pipelines and deploying models at scale while minimizing hallucinations.

Monitoring and Maintenance: Continuous monitoring to detect performance issues and ensure reliability.

Talent and Expertise: Recruiting and retaining skilled professionals in AI, machine learning, and cloud computing.



SPEAKER SPOTLIGHT
Galina Sagan
Prinicpal
Hitachi Ventures





WHY DO YOU THINK IT IS SO IMPORTANT TO ATTEND AND PARTICIPATE IN THE AI HARDWARE & EDGE AI SUMMIT EUROPE?

Gain insights into the latest advancements in AI hardware, edge AI infrastructure, software infrastructure, and ML applications:

- LLM Model design architecture
- Cost and energy efficient training and inference methods for LLMs
- Improving accuracy, dealing with hallucinations in enterprise grade LLMs etc

Connect with European AI leaders deploying solutions in manufacturing, telecommunications, finance, and healthcare and several other industries.

Discover emerging startups and talent in Europe's Al space, fostering potential partnerships or investment opportunities.

Focus on edge Al and efficient deployments, addressing challenges and solutions for deploying Al models on resource-constrained edge devices.

Explore real-world use cases through case studies and presentations from enterprises, offering insights into successful Al and ML deployments.



SPEAKER SPOTLIGHT Galina Sagan Prinicpal

HITACHI Inspire the Next



WHY DO YOU THINK IT IS SO IMPORTANT TO ATTEND AND PARTICIPATE IN THE AI HARDWARE & EDGE AI **SUMMIT EUROPE?**

Efficient Deployment of Large Language Models and Generative AI:

- Optimizing hardware and software for edge deployment
- Techniques for compression and hardware-aware optimization

Advancements in Edge AI Hardware and Systems:

- Latest developments in specialized AI chips and neuromorphic computing
- Innovations for energy efficiency and thermal constraints

Scaling Edge Al Infrastructure for Industrial Use Cases

- Challenges and solutions for industrial settings
- Real-world case studies in manufacturing, automotive, etc

Software Infrastructure and Tooling for Edge Al:

- Advances in frameworks and tools for model optimization.
- Techniques for efficient quantization and compilation

Emerging Al Hardware Startups and Disruptive Technologies:

- Presentations from innovative startups on specialized chips, neuromorphic computing, etc
- Insights into potential impacts on edge AI