



Organized by **Mehta Family
School of
Data Science
& Artificial
Intelligence**

Indian Institute of Technology
Guwahati,
Guwahati-781039, India

The Future of Multimodal AI Applications

**Move beyond text-based paradigms and explore the frontier of Artificial Intelligence:
Multimodal applications that see, hear, and interact with the world in real-time**



Stefania Druga, PhD

Research Scientist,
Google DeepMind

This talk will explore the practical implementation and potential of systems integrating and synthesizing information from diverse data streams - including live input from webcams, audio feeds, video, mobile sensors, and bespoke hardware. Stefania Druga (Google DeepMind) will provide a researcher's perspective, diving into the technical challenges - from data fusion and latency reduction to context modeling and robust interaction design - inherent in building AI that leverages a richer understanding of the physical world. Expect compelling live demonstrations showcasing interactive AI systems designed to understand context, anticipate user needs, and respond dynamically through multi-sensory feedback loops.

About the Speaker

Stefania Druga is a Research Scientist at Google DeepMind, where she works on novel multimodal AI applications. She has a master's degree from MIT, PhD from UW and has been doing research on AI education since 2015. During graduate school, she built the first open-source platform for K12 AI Education - Cognimates. When she is not coding and writing papers, she enjoys trail running, yoga, and riding her bicycle.

Website: <https://stefania11.github.io/>

**Venue: Online
Time: 01:30 PM IST
Date: Monday 12-May-25**