

# KEYNOTE TALK

Monday, November 17, 2025 at 9:00 am

## The Curious Case of Foundational and VLM Models

Leonid Sigal

University of British Columbia

**Abstract:** The capabilities and the use of foundational (FM) and vision-language (VLM) models in computer vision have exploded over the past few years. This has led to a broad paradigm shift in the field. In this talk I will focus on the recent work from my group that navigates this quickly evolving research landscape. Addressing challenges such as building foundational models with better generalization, increasing their context length, adopting them to ever evolving task landscape and routing information among them for more complex reasoning visual problems. I will also discuss some curious benefits and challenges of working with such models, including emergent (localization) capabilities and in-consistency in their responses.



**Speaker Bio-Sketch:** Prof. Leonid Sigal is a Professor at the University of British Columbia (UBC). He was appointed CIFAR AI Chair at the Vector Institute in 2019 and an NSERC Tier 2 Canada Research Chair in Computer Vision and Machine Learning in 2018. Prior to this, he was a Senior Research Scientist, and a group lead, at Disney Research. He completed his Ph.D at Brown University in 2008; received his B.Sc. degrees in Computer Science and Mathematics from Boston University in 1999, his M.A. from Boston University in 1999, and his M.S. from Brown University in 2003. Leonid's research interests lie in the areas of computer vision, machine learning, and computer graphics; with the emphasis on approaches for visual and multi-modal representation learning, recognition, understanding and generative modeling. He has won a number of research awards, including Killam Accelerator Fellowship in 2021 and has published over 100 papers in

venues such as CVPR, ICCV, ECCV, NeurIPS, ICLR, and Siggraph.