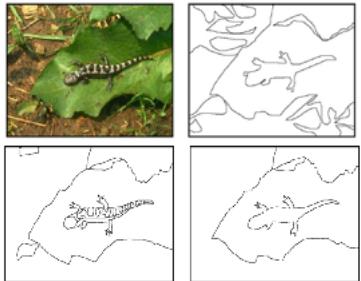


The third-person Web perceptual experience



BSD (2001)

A curated “disembodied” moment in time
from a spectator’s perspective



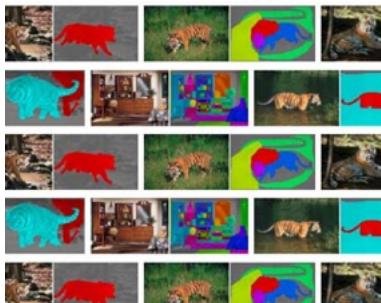
Caltech 101 (2004), Caltech 256 (2006)



PASCAL (2007-12)



ActivityNet (2015)



LabelMe (2007)



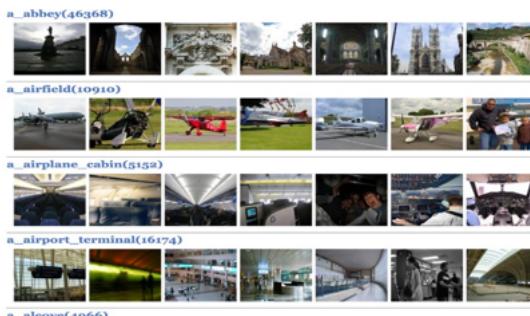
ImageNet (2009)



SUN (2010)



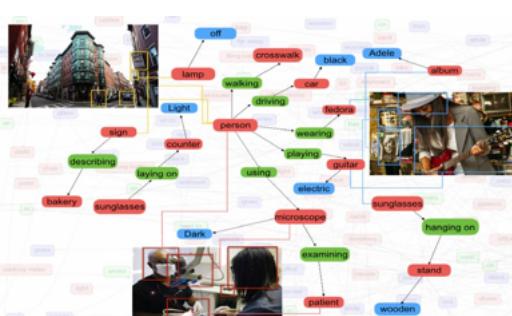
Kinetics (2017)



Places (2014)



MS COCO (2014)



Visual Genome (2016)



AVA (2018)

First-person “egocentric” perceptual experience

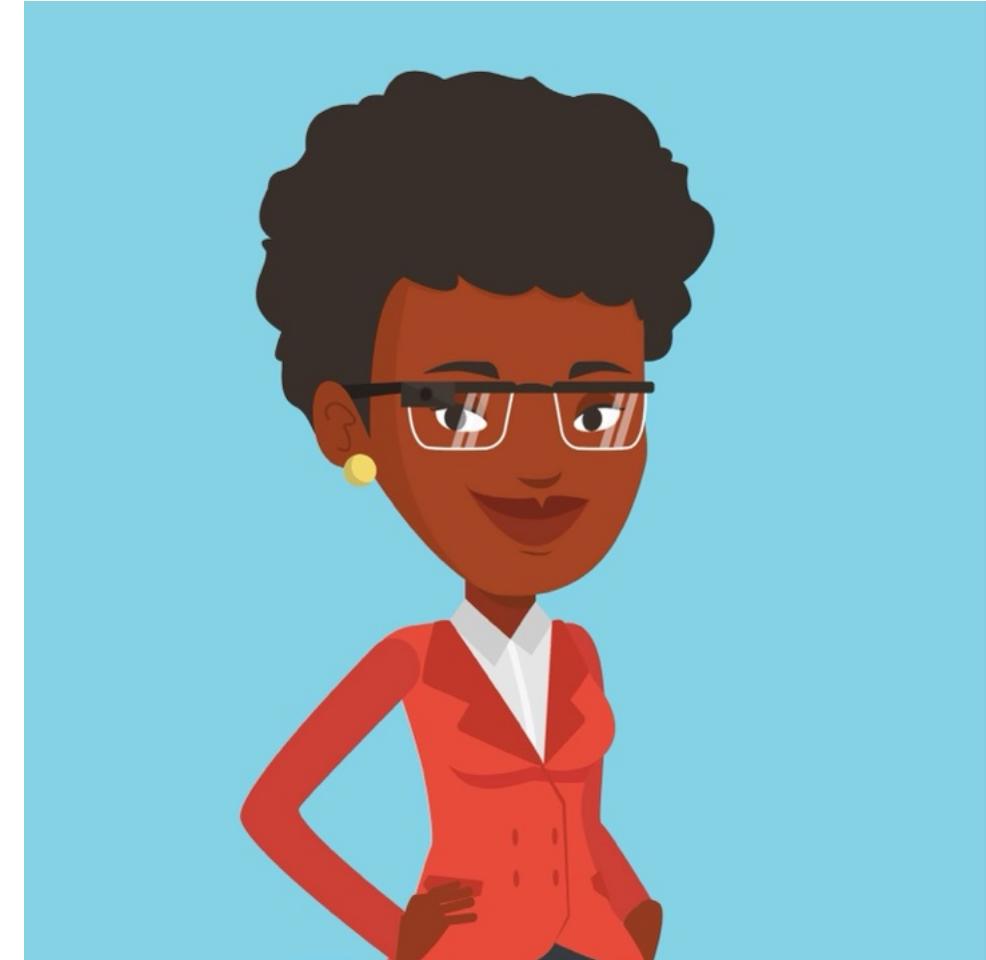
Uncurated long-form video stream driven by



Why egocentric vision?

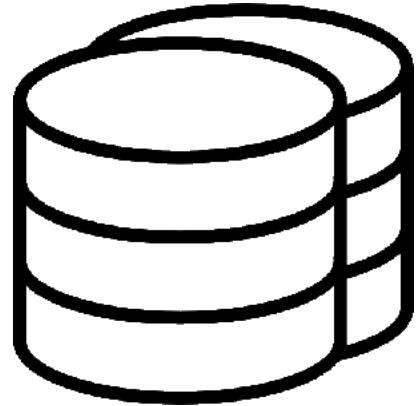


Robot learning



Augmented reality

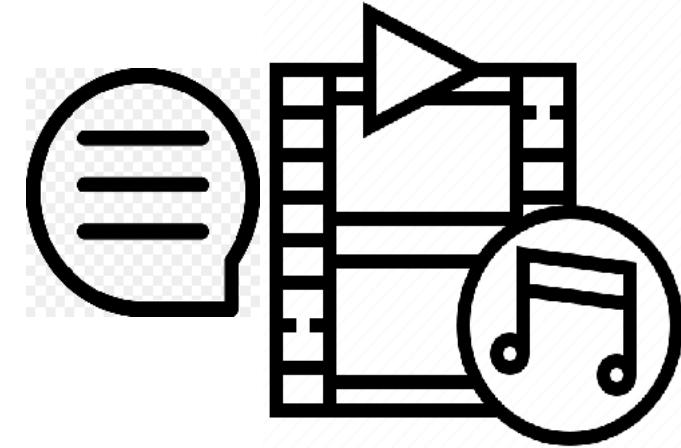
First-person vision: examples in research



Data foundations



Affordances and manipulation



Multimodal learning

Existing first-person video datasets

Inspire our effort, but call for greater scale, content, diversity



EPIC Kitchens

Damen et al. 2020

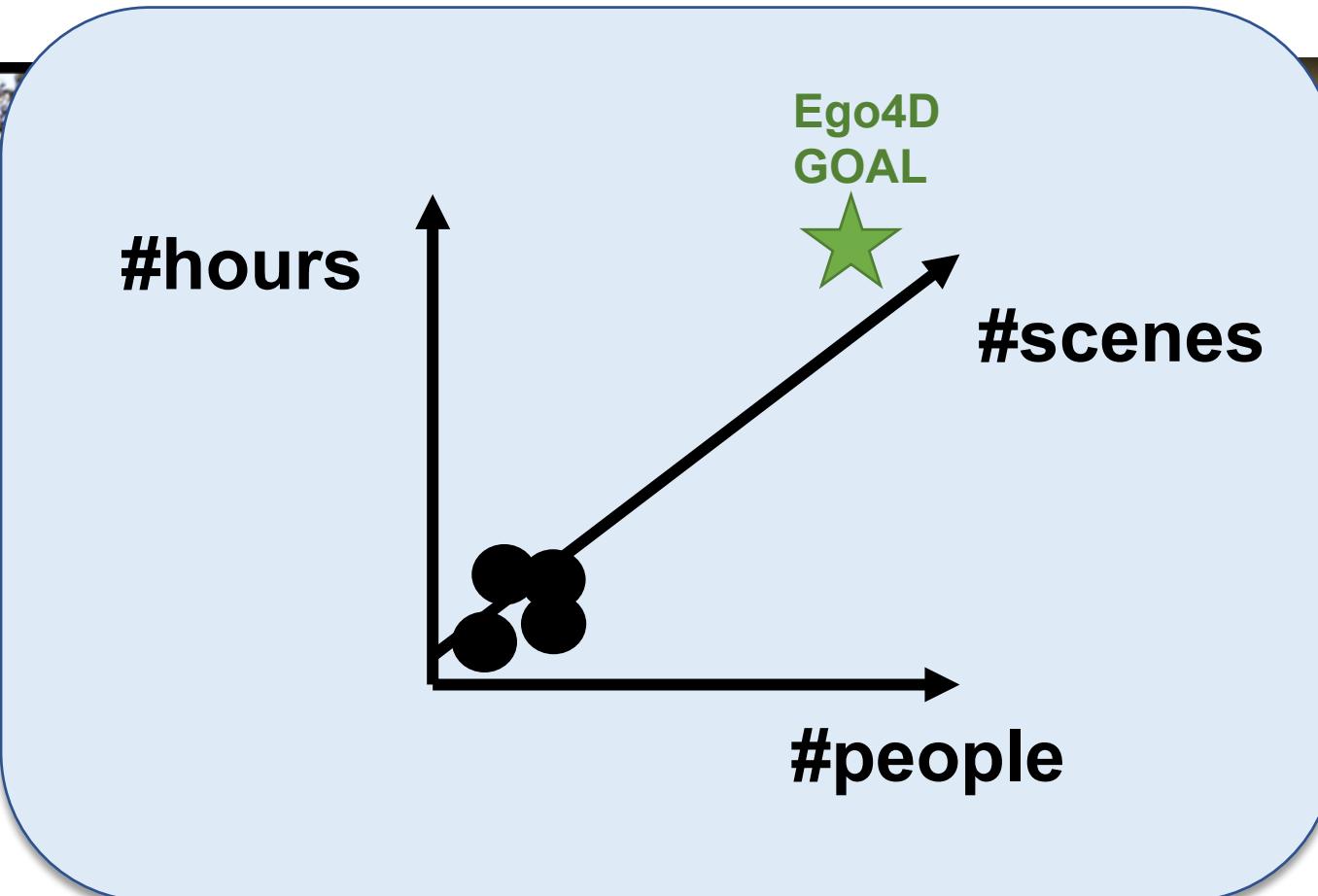
45 people, 100 hrs
kitchens only



UT Ego

Lee

**4 pe
daily**



Charades-Ego

Sigurdsson 2018

71 people, 34 hrs
indoor



Existing first-person video datasets

 # Participants

Hours

EPIC-Kitchens-100



EPIC-Kitchens-100

Ego4D: A multi-sensor scale video dataset

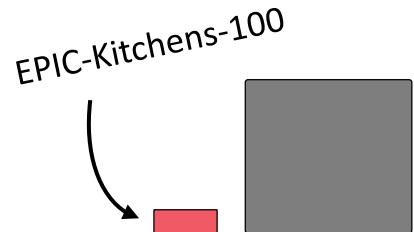


3,670 hours of in-the-wild daily life activity

931 participants from 74 worldwide locations

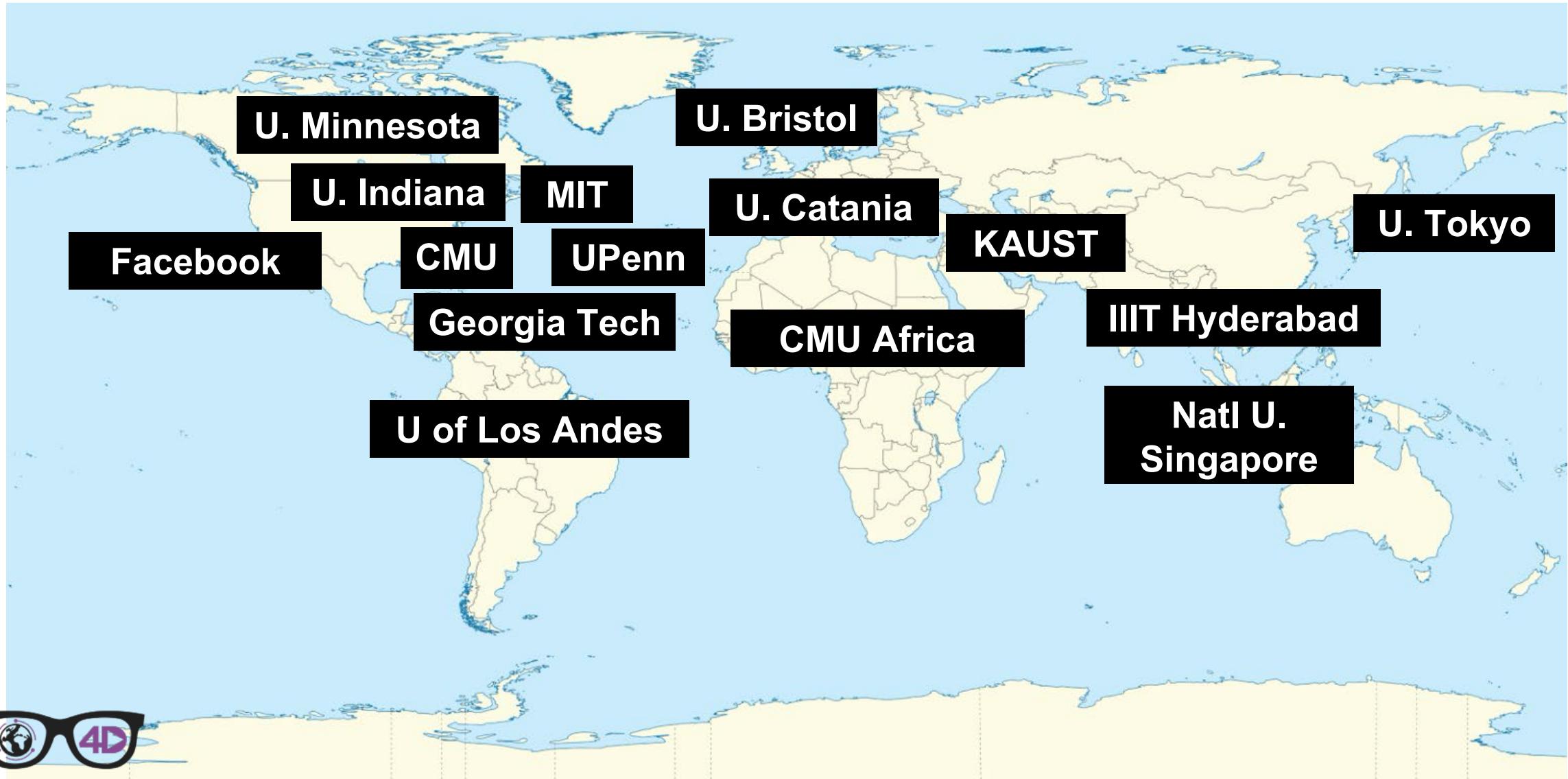
Multimodal: audio, 3D scans, IMU, stereo, multi-camera

Benchmark tasks to catalyze research

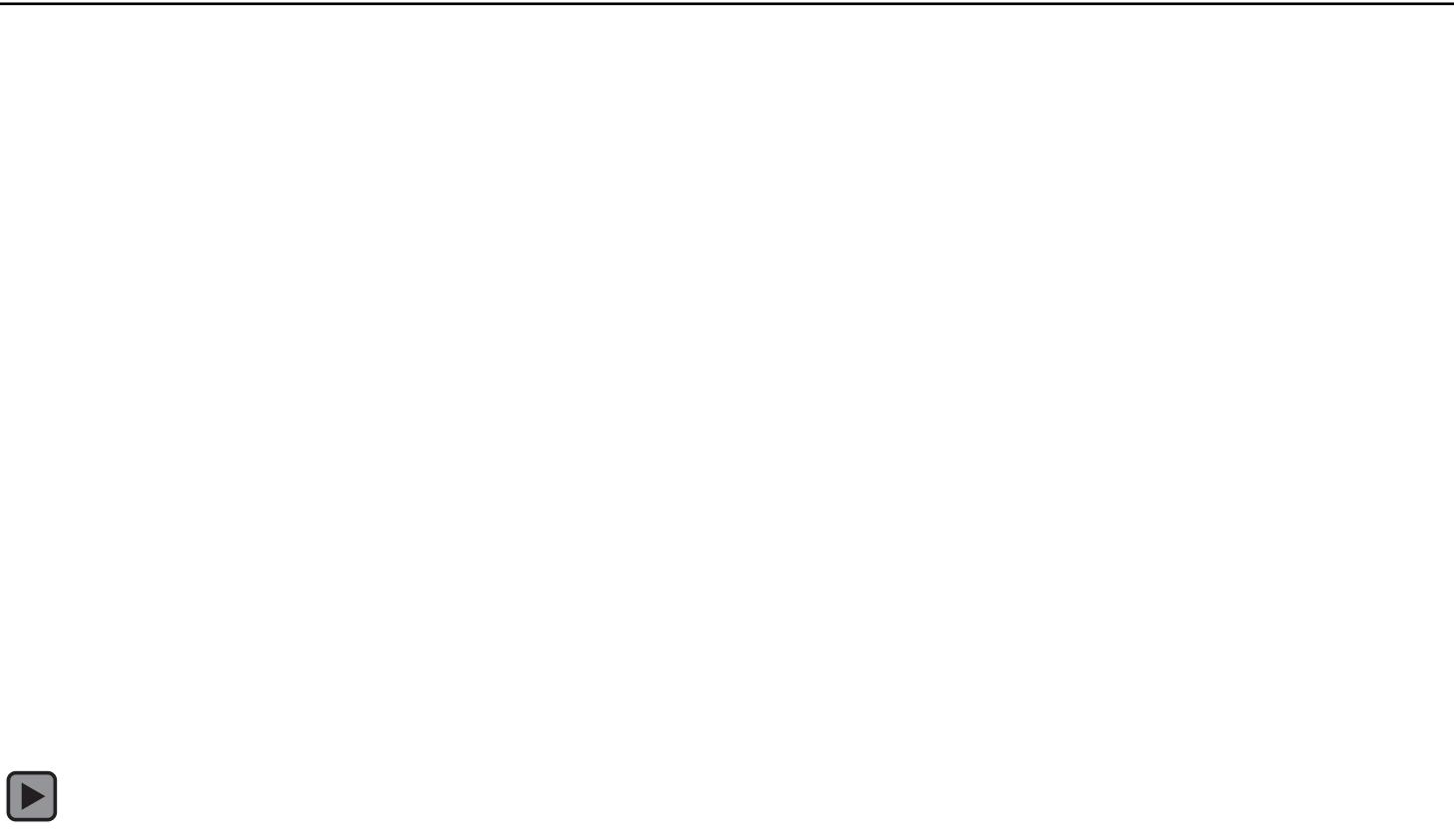


Ego4D consortium

Towards diverse geographic coverage



Ego4D: everyday activity around the world



3,670 hours of in-the-wild
daily life activity

931 participants from 74
worldwide locations

Benchmark tasks to
catalyze research

Multimodal: audio, 3D scans,
IMU, stereo, multi-camera, text
narrations



Ego4D benchmark suite

Past



Episodic Memory
“where is my X?”

Present



Hands & Objects
“what am I doing and how?”

Present



Audio-visual Diarization
“who said what when?”



Social Interaction
“who is attending to whom?”

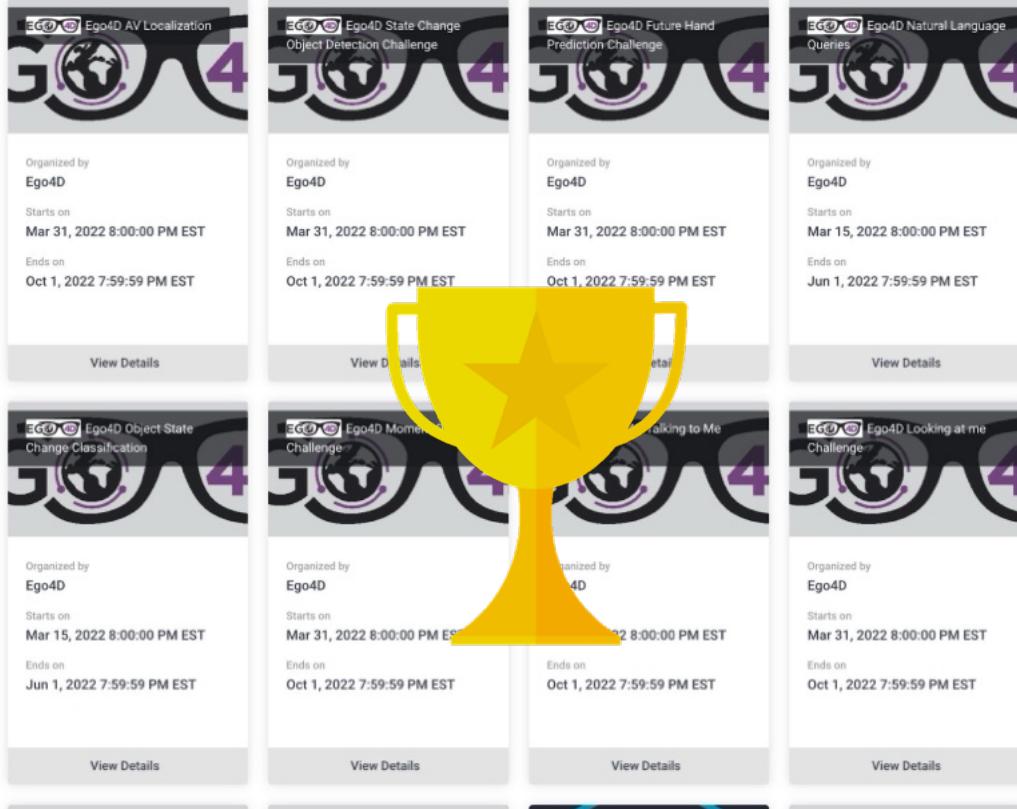
Future



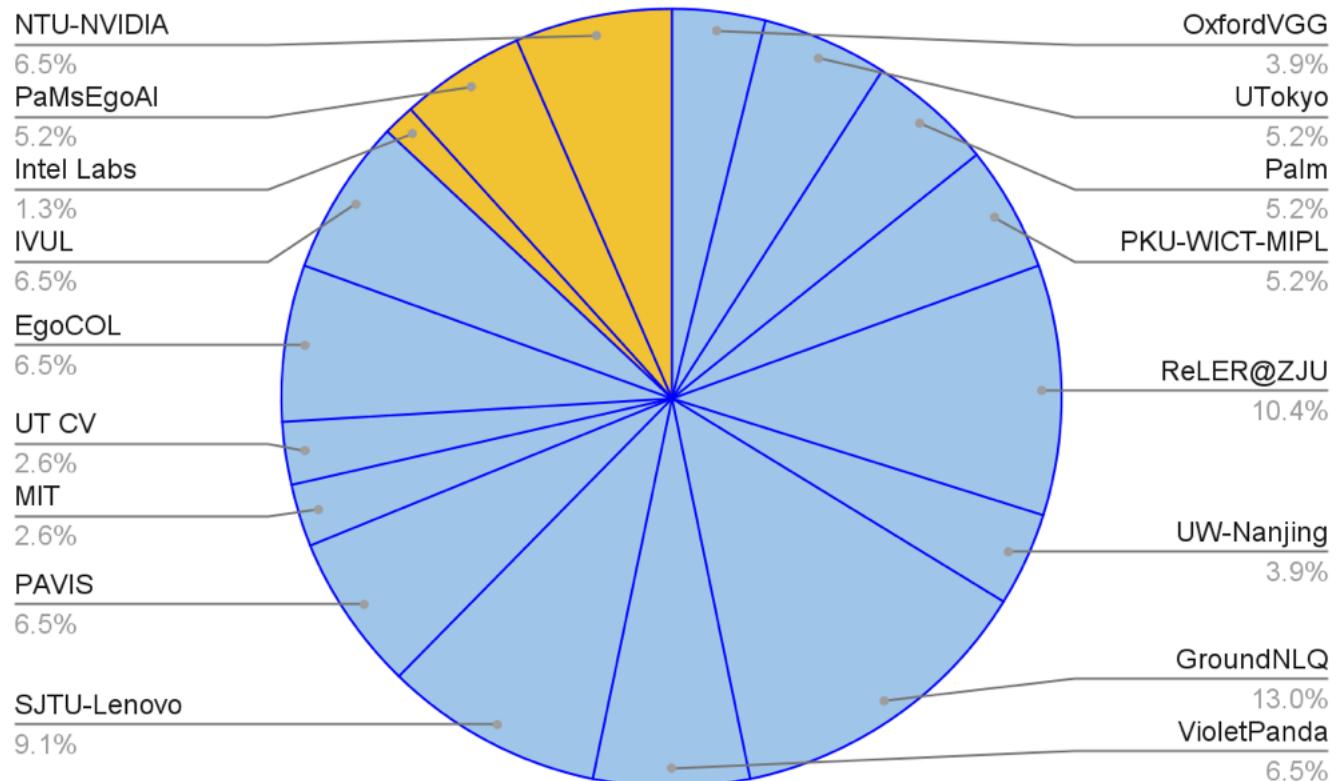
Forecasting
“what will I do next?”

Ego4D challenges

Broad and growing participation at CVPR 2022, ECCV 2022, CVPR 2023



Hosted on EvalAI

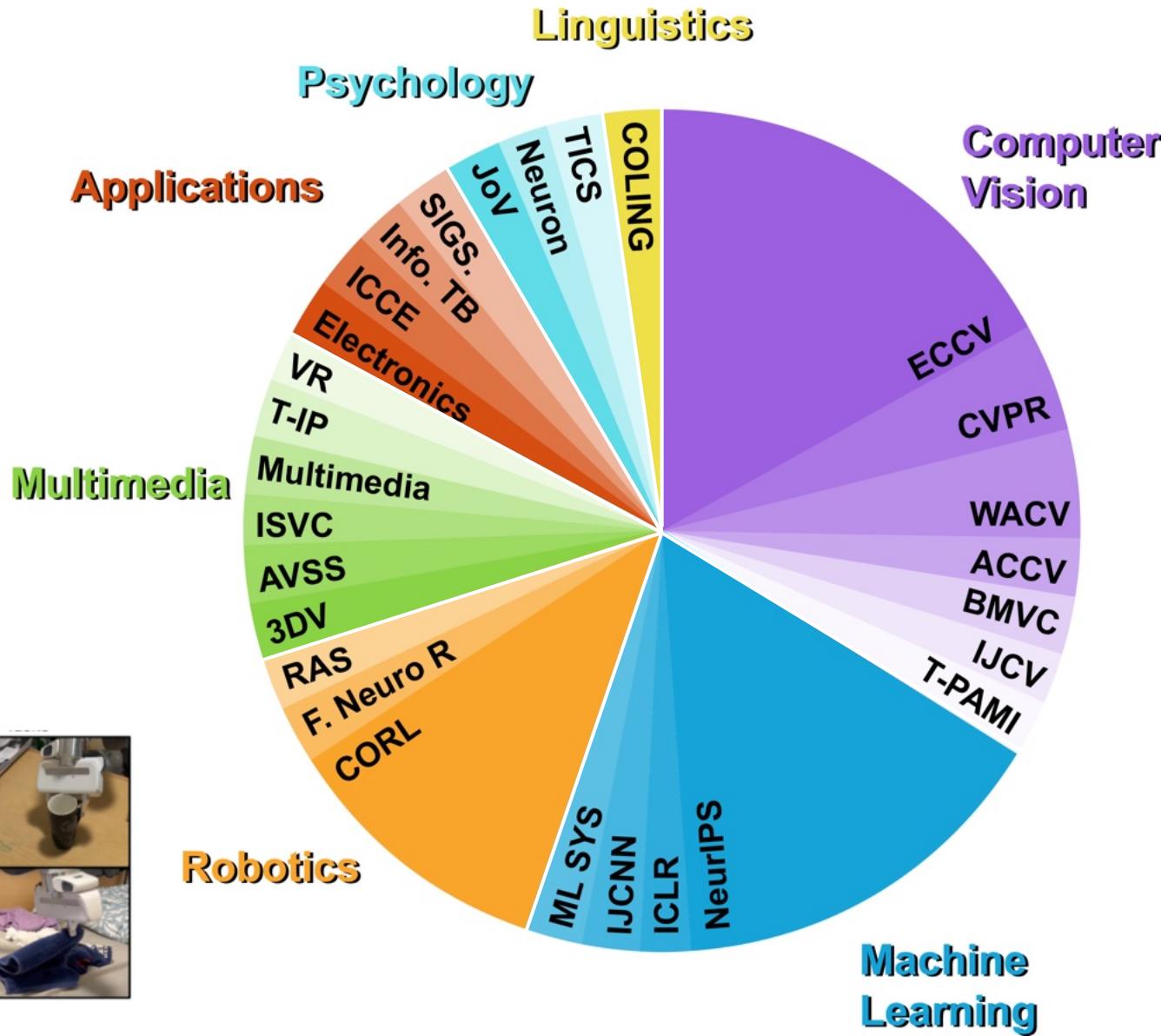


Industry vs Academic teams

<https://ego4d-data.org/docs/challenge/>



Ego4D is enabling new research in the broader community



Papers citing Ego4D
as of 4/2023

Source: Google Scholar
Slide credit: David Crandall