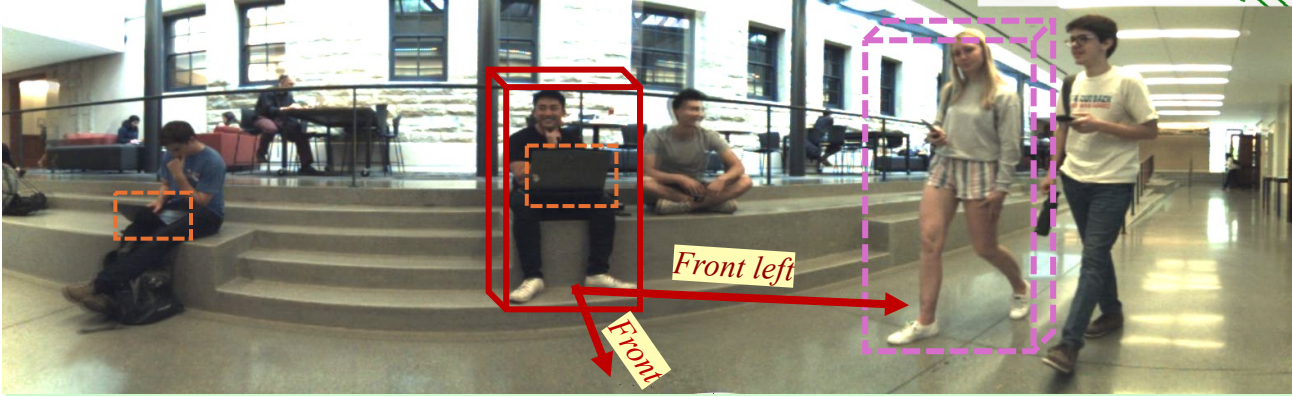
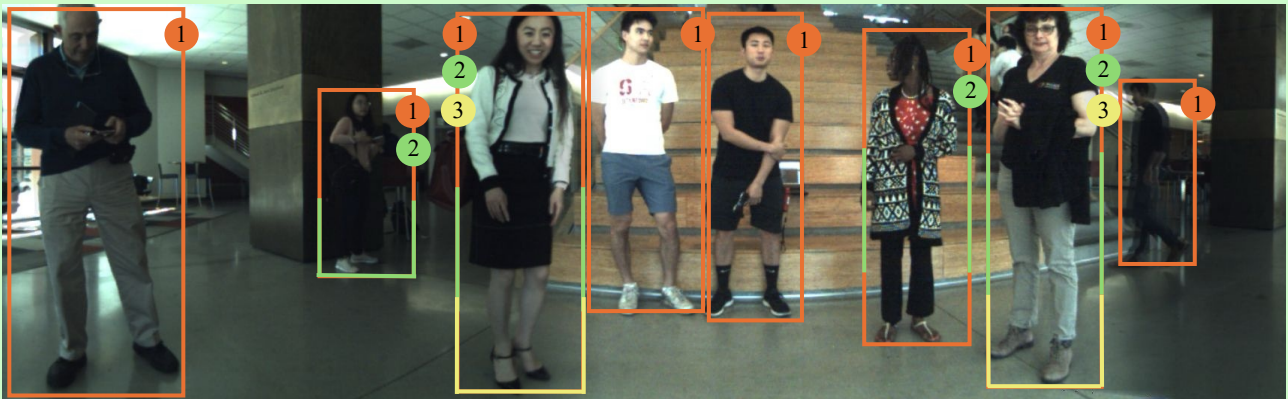


# JRDB-Reasoning: A Difficulty-Graded Benchmark for Visual Reasoning in Robotics

(a) -Find a *male Asian* person who is *sitting* and *working with a laptop*, with a *female* positioned to his *left front* at a *close distance*.

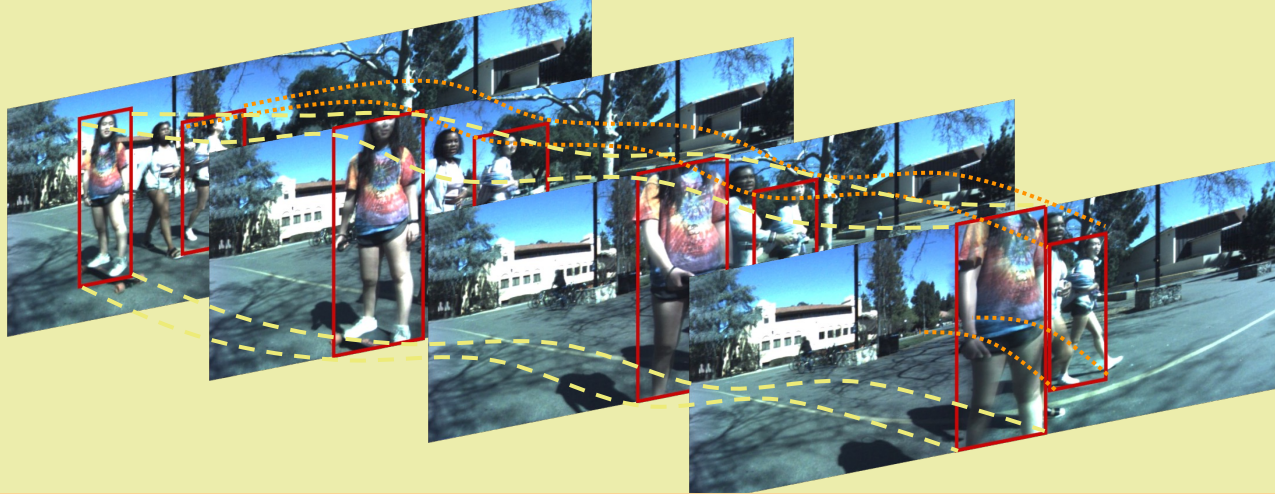


(b) -Count the number of *female* persons who *are looking at the robot (me)*.



Options: 9, 6, 4, 0, 1, 3, 2, 5, 8, 11, 7, 10, 14, 13.      Answer : 2

(c) -Track *Asian* persons who are *walking* during 45 to 90 seconds.



(d) -What is the geometric location of the bicycle relative to me (the robot) between 31.8 and 33.07 seconds?



Options: [(right, front, left),(back ,left, fornt left),..... ]      Answer : left, front, right

(e) -Locate/track a *male* person who is *standing* alone, *carrying a bag*, and a *female* person is located *to his back*.



Human Object Interaction:    ●-sitting    ●-holding    ●-touching    ●-carrying    ●-standing