# **Multi-modal Situated Question Answering**

### **Situation:**

 $(s^{loc}, s^{rot})$ 

### **Situation description:**

I'm sitting on a blue bed. A on my 2 o'clock.





## **Question:**

What is the spatial relationship between the



and the

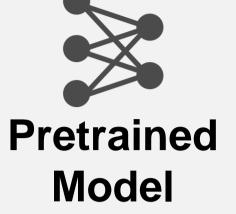


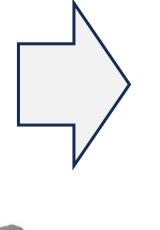
**Answer:** The wardrobe is on the left of the table.

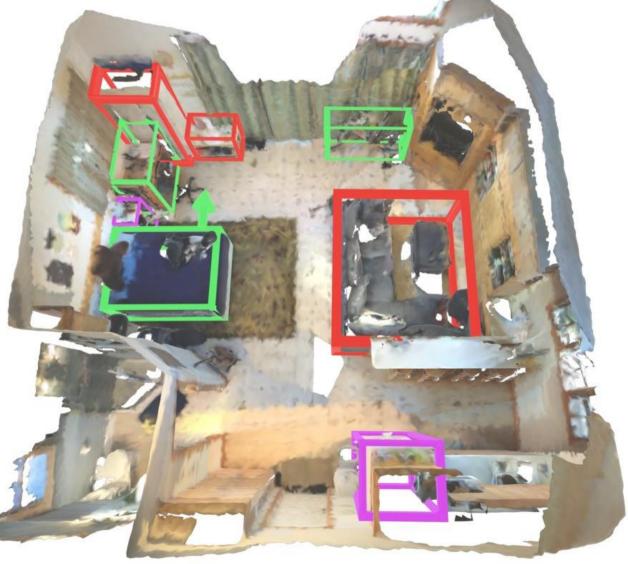
Question: Describe the sofa on my right.

**Answer:** The L- shape sofa is blue with the soft fabric material. There are some pillows on the sofa.









# Multi-modal Situated **Next-step Navigation**

#### **Situation:**

 $(s^{loc}, s^{rot})$ 

## **Situation description:**

I'm sitting on a blue bed. A is on my left.



### **Instruction:**

I want to get some food from the refrigerator. What action should I take next step?

Action Response: Turn right.

#### **Instruction:**

I want put a cup on the white nightstand. The nightstand is near a desk. What action should I take next step?

Action Response: Turn left.