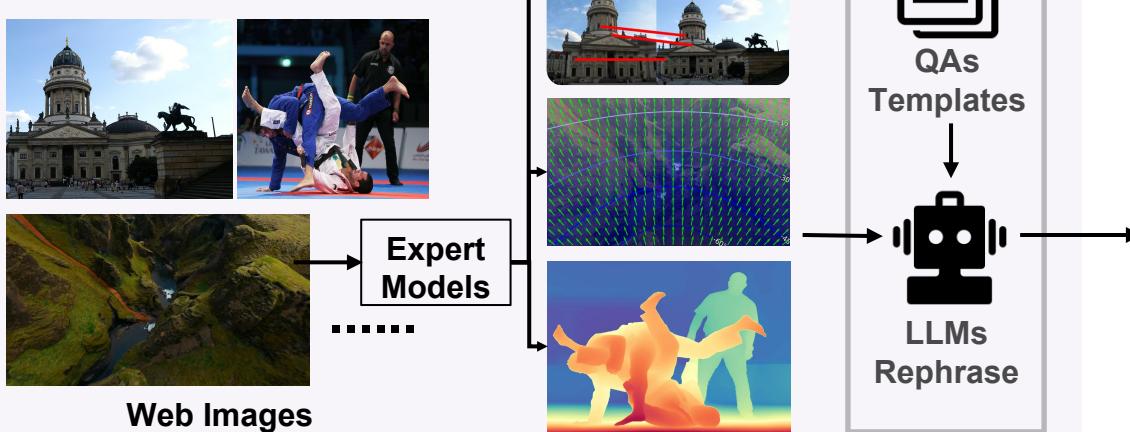
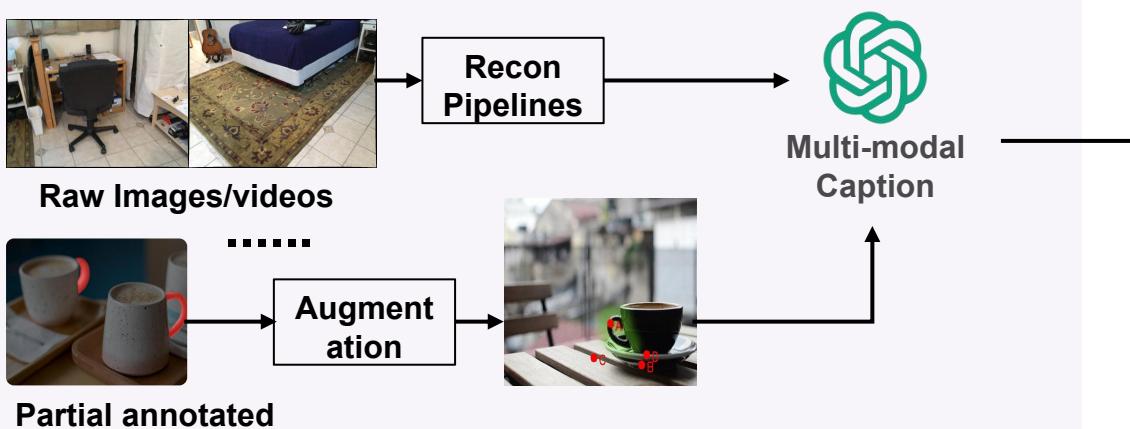


Data Engine

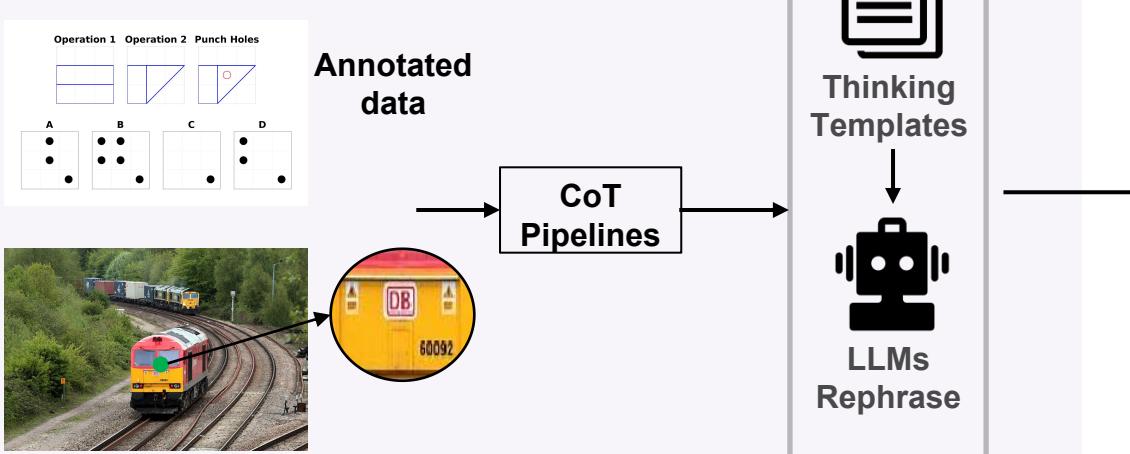
L1 Data Engine



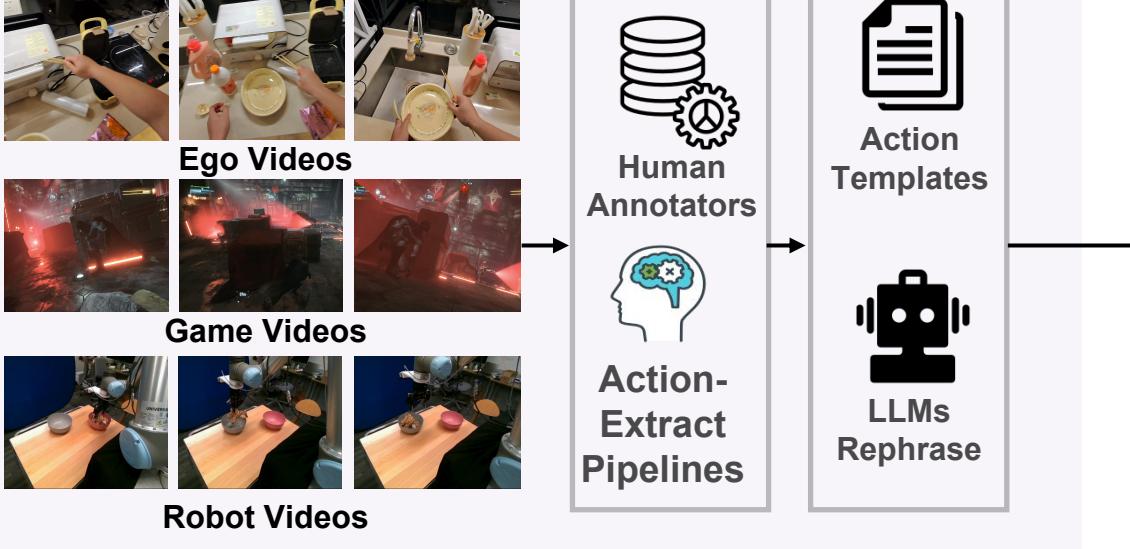
L2 Data Engine



L3 Data Engine

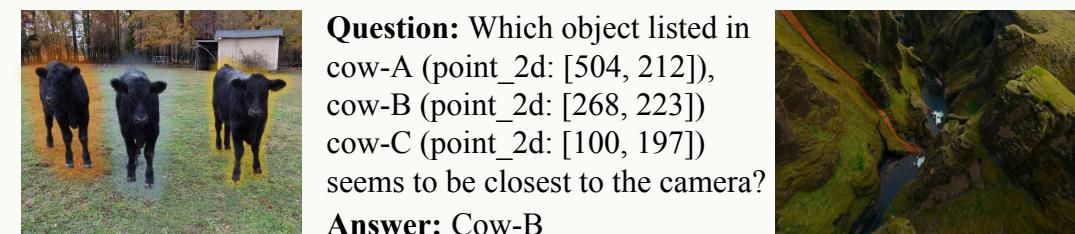


L4 Data Engine



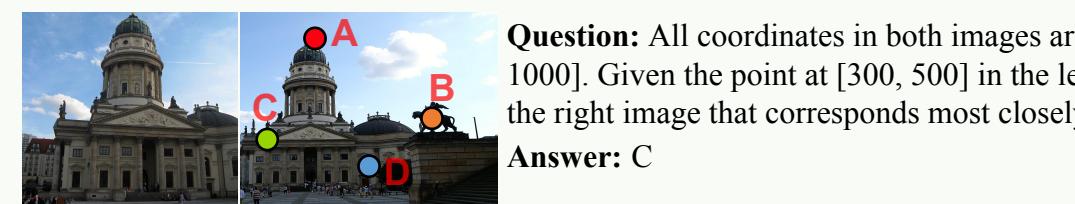
Instantiated Samples

Samples for Basic Perception



Question: Please estimate the gravity.

Answer: Roll: 1.10° ($\pm 20.54^\circ$) Pitch: -38.71° ($\pm 9.22^\circ$)



Samples for Mental Mapping

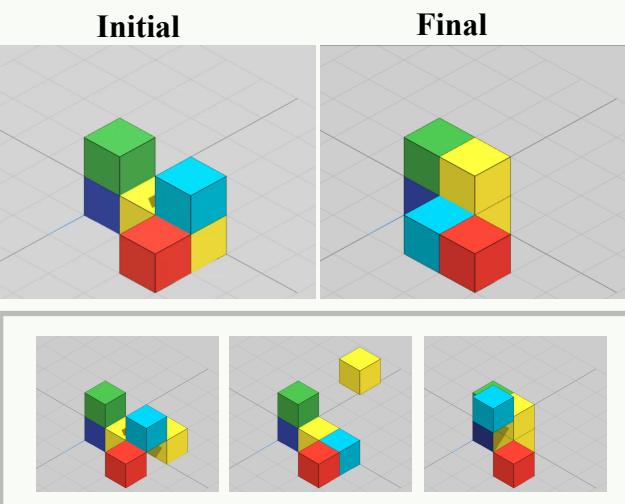


Question: Please describe the relation between these two buildings.

Answer: They look very similar in structure and layout.



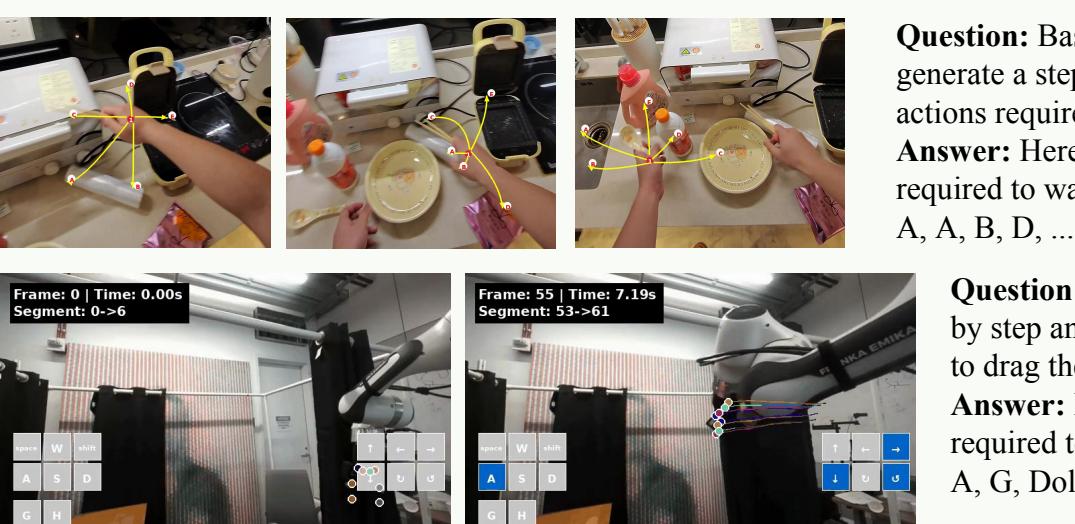
Samples for Mental Simulation



Question: Please provide the transformation sequence for changing the multidimensional dataset stack from the initial state to the final state shown in the figure.

Answer: Here is the sequence to transform.

Samples for Spatial Agents



Question: Please control the gripper step by step and generate a sequence of actions to drag the curtain from right to left.

Answer: Here is the sequence of actions required to wash the dish:
A, G, Dolly in, ...