

01 Source data

BEHAVIOR-1K demos

2025 + 2026 challenge releases



Manipulation



Navigation

Recorded observations

RGB + instance segmentation
Robot state + actions

Simulator geometry

Object poses + extents
Camera calibration
Replay / restore as needed

02 Metadata + task curation

Extract spatial metadata

Visible objects · 2D boxes / masks
3D relations · reference frames

Direction · distance · size

Curate task instances

SELECTION CRITERIA

Spatial relation determines the target;
manipulation has confusable candidates.

CURATION STEPS

1. Select skills and trim task clips
2. Define relation-changing prompts
3. Assign target objects / goal regions
4. Curate navigation context tours

OUTPUT

Scene state + instruction + target
Task clips + navigation context

03 Generate and validate

General VQA

GENERATE Build MCQs; balance categories

CHECK Relation margins + option visibility

Task-grounded VQA

GENERATE Build task MCQs; balance variants

CHECK Question-relevant objects visible

Grounding labels / IR

GENERATE Compute target points / regions

CHECK Region visible; oracle / random checks

Action data + eval states

GENERATE Export task clips + saved states

CHECK Split novelty + valid, visible states

04 Benchmark suites

General spatial VQA

Single-frame + multi-frame

Task-grounded VQA

Per-task understanding

Grounding suite + IR

Pointing evaluation
Oracle policy inputs

VLA datasets

14 manipulation tasks
10 navigation tasks