

REAL-WORLD ROBOT DATA

真实世界经验， 如何成为机器人可学习的数据

从一条真机轨迹，到可训练、可评测、可持续回流的数据基础设施

先把“机器人数据”问具体

- 01 一条可训练的机器人轨迹，究竟必须包含什么？
- 02 我们如何证明它在时间、空间与任务结果上可信？
- 03 它如何进入训练、评测，并把失败再次送回数据队列？

这不是“拍了多少小时视频”的问题，而是：能否还原一次完整、可对齐、可追责的物理事件。

同样叫“数据”，解决的是不同问题

1M+ Open X-Embodiment 22 种机器人形态，强调跨数据集、跨具身的规模与泛化。	76k / 350h DROID 真实环境、分布式采集，强调场景与操作者的多样性。	IN-THE-WILD UMI 便携式人类示范采集，探索低门槛采集与真机策略迁移。	MIXTURE GR00T N1 将人类视频、真机、仿真与合成数据放进异构训练混合物。
--	--	--	--

规模、多样性、动作可观测性与物理可信度，是四个不同的工程轴。

视频不是轨迹，轨迹也不是唯一答案

数据形态	动作可观测	接触与本体	优势	主要缺口
互联网 / 人类视频	弱	弱	规模大、语义丰富	缺少可执行动作
UMI / 第一视角示范	中	部分	便携、可进真实场景	需完成具身映射
仿真 / 世界模型	强	可控近似	低风险、覆盖长尾	存在现实差异
真机遥操作轨迹	强	强	动作与物理结果同源	成本与治理复杂

判断原则：不存在普适排序。要看任务是否依赖精细接触、动作可逆性、具身约束与部署环境。



世界模型扩展“可能发生什么”

Cosmos：可生成、可预测的物理 AI 世界

通过世界基础模型与数据处理框架，服务物理 AI 的合成数据生成、预测和评估。

公开来源：NVIDIA Research, 2025

Genie 2：由图像提示生成可交互环境

探索大规模世界模型对多样环境与交互未来的生成能力。

公开来源：Google DeepMind, 2024

覆盖

生成更多环境、初始状态与失败分支。

校准

真机数据用于确认感知、动作与接触偏差。

验证

部署成功率仍需回到真实机器人上测量。

我们的工程判断：生成经验负责扩覆盖，真机轨迹负责校准与验收，二者不是替代关系。

一条可训练轨迹，是七类信号的共同时间线

```
{
  "trajectory_id": "demo_7f2a",
  "clock": "ptp_hw_01",
  "frames": ["head_rgb", "wrist_l", "wrist_r"],
  "state": { "joint", "eef", "gripper" },
  "action": { "commanded", "executed" },
  "contact": { "force", "event" },
  "calibration_id": "cal_2026_04",
  "outcome": "recovered_success"
}
```

不含客户、场地、操作者等可识别信息；原始数据与发布样例采用不同标识域。

同一事件，必须能回答七个问题



关键：时间戳和标定版本不是附属元数据，它们决定这些信号能否组成同一个监督样本。

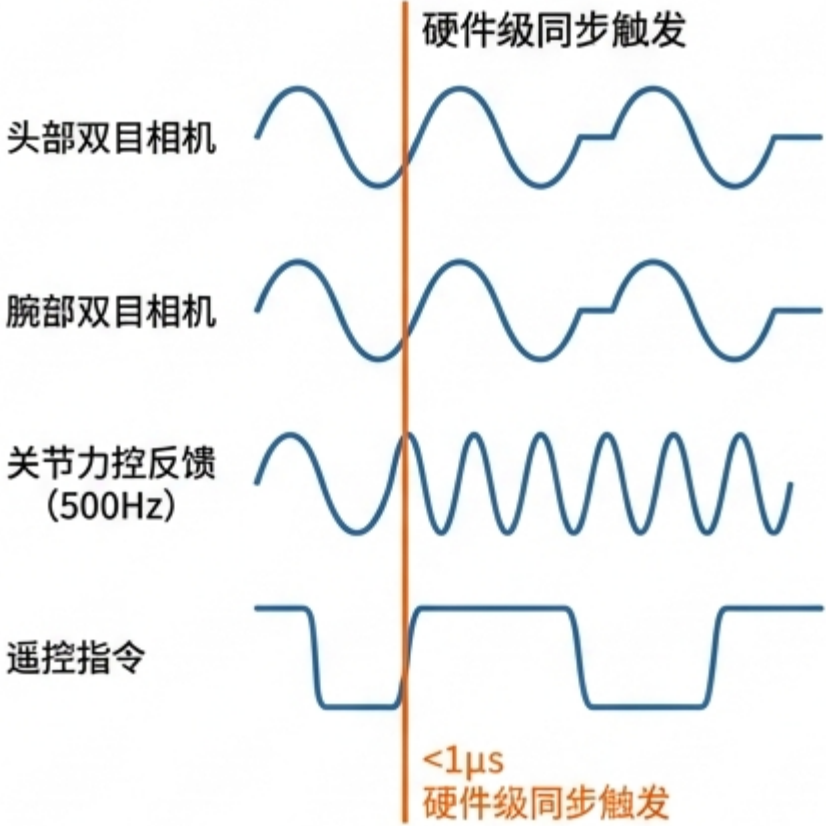
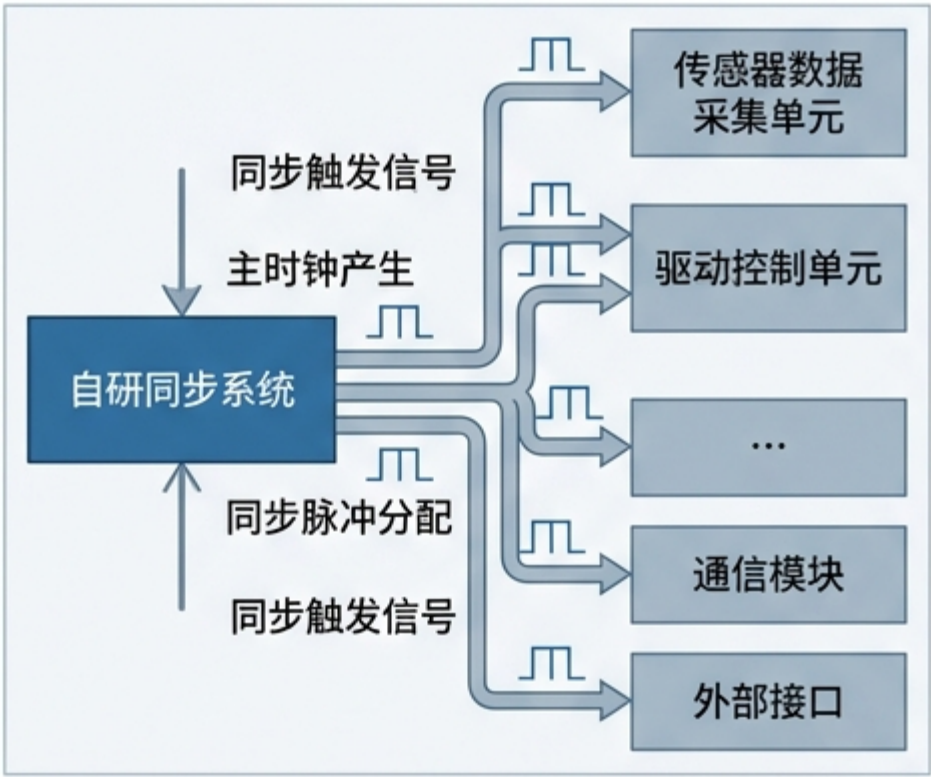
错一帧，动作监督就可能指向错误的世界

<1μs

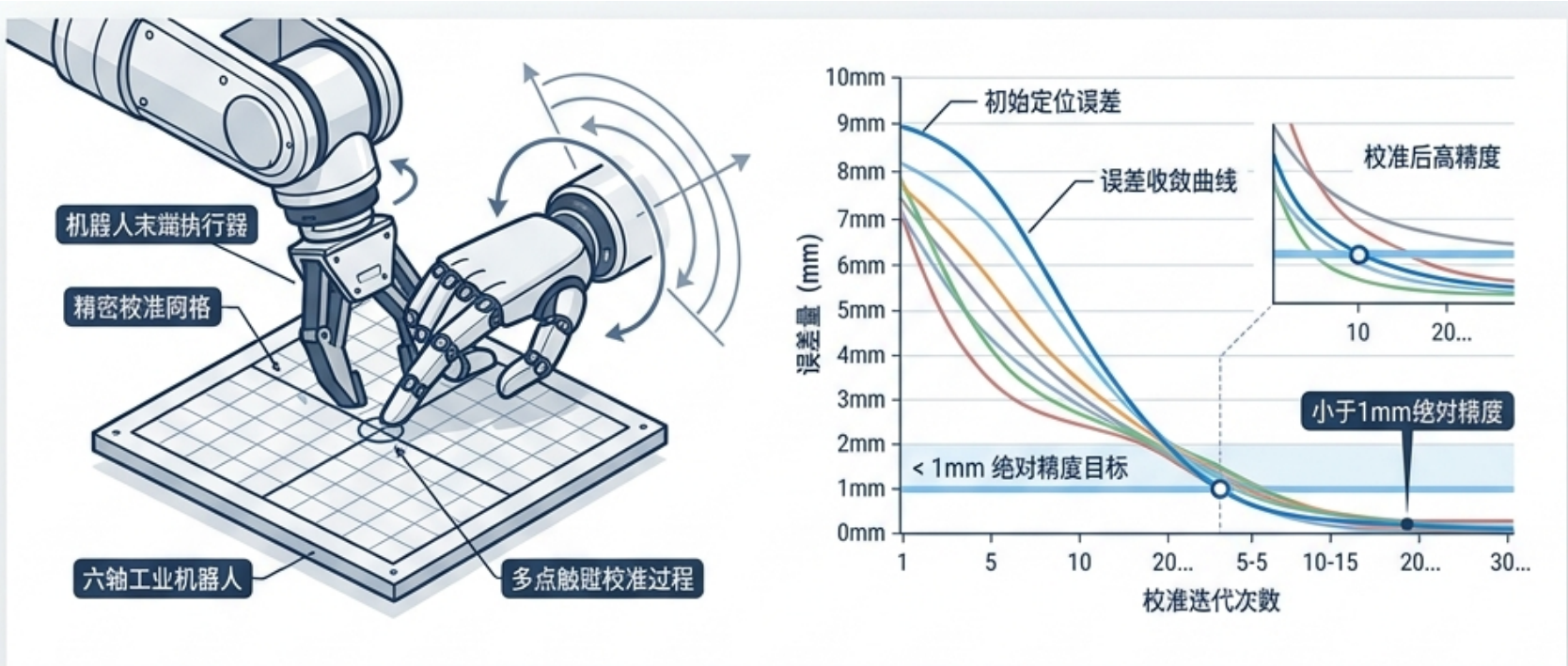
时间同步精度

6 路相机硬件同步，让视觉、遥控指令与 500Hz 关节力控反馈落在可对齐的时间基准上。

测试口径：同步控制器主时钟到各硬件触发端的通道间偏差。数值沿用 BP3 已披露工程口径；测试仪器、负载和样本量尚未形成公开复现报告。



知道“何时”，还要知道“在哪个坐标系”



<1mm

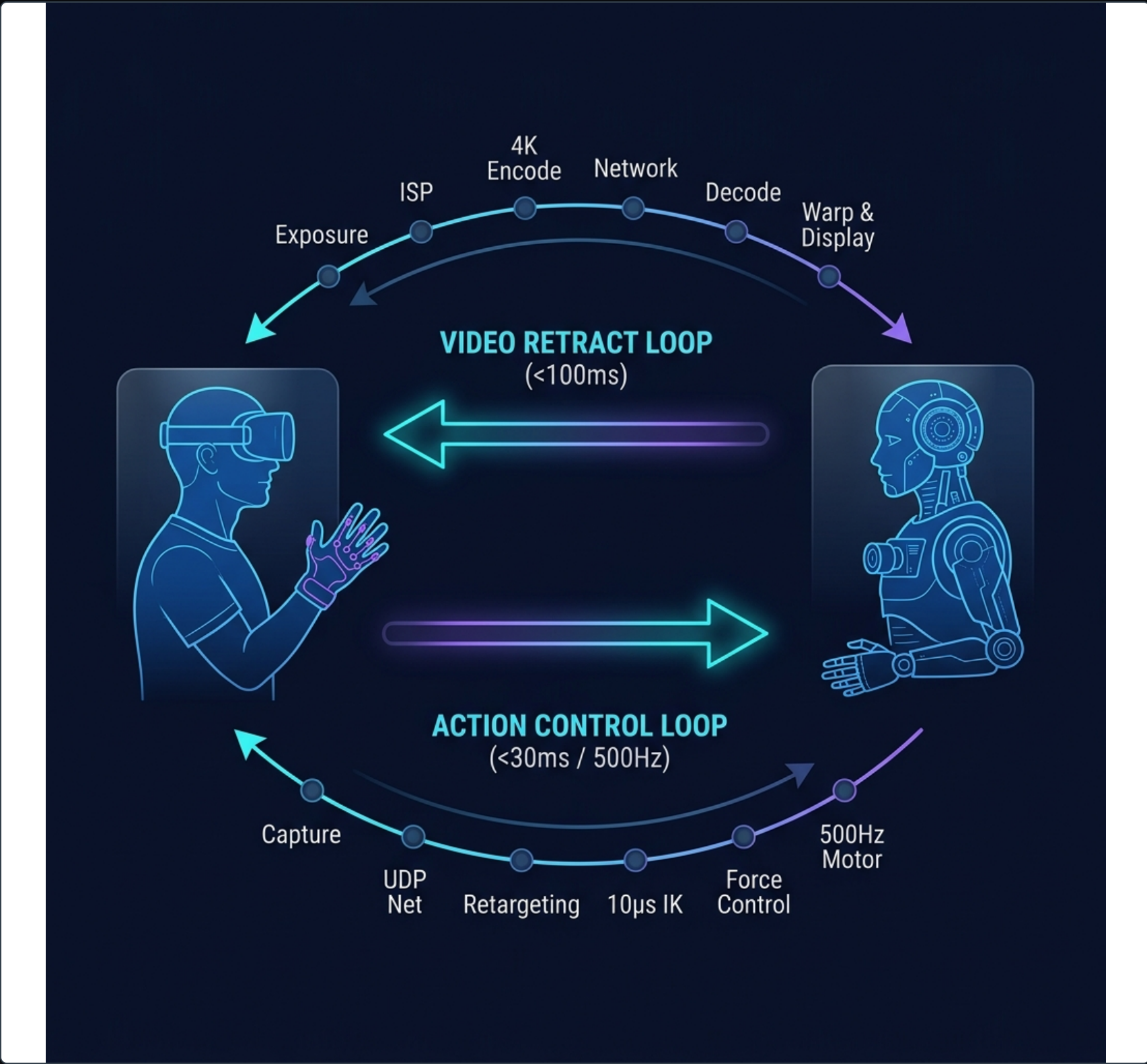
空间定位精度

把相机坐标、机器人基座坐标、末端执行器坐标与任务物体坐标连成可追溯的变换链。

动作表示是否可复现，取决于坐标系定义、标定版本和漂移检测，而不只是模型输出格式。

测试口径：末端执行器对精密校准网格目标点执行多点触碰，比较目标点与校准后实际触点的绝对残差。数值沿用 BP3 图示口径，未作独立审计声明。

遥操作不是一个延迟，而是两个闭环



视频文字说明：操作者通过手持控制器输入意图，机器人端执行动作并回传多视角画面。观察重点是动作启动、视觉反馈与接触后的修正是否处于同一闭环，而非只看单次演示是否成功。

图示中的链路边界：曝光/编码/网络/解码/显示属于视频反馈环；采集/网络/重定向/逆解/力控/电机属于动作控制环。

把“低延迟”拆成可复测的工程量

指标	测量边界	测试口径 / 当前公开限制	对数据的影响
<100ms 灵御内部实测	图像传输延迟	从相机曝光链路开始到操作端显示完成；包含编码、网络、解码与显示。公开版暂不列网络条件分布。	决定操作者感知与纠错时机
<30ms 灵御内部实测	本地控制延迟	从本地动作输入采集到机器人控制接口接收；不与图像链路合并计算。	决定意图和执行动作的时序一致性
10μs 灵御内部实测	底层控制周期	逆解 / 底层控制计算周期的 BP3 工程配置口径；不是端到端遥操作延迟。	决定高频反馈可用性
0–10N 灵御内部实测	关节力控范围	当前系统公开的力控工作范围；机器人型号、关节与负载条件需在复现报告中逐项列出。	支持接触事件与安全边界记录
250+ 灵御内部实测	条 / 小时	按已验收轨迹或原子动作计数的小时吞吐；任务复杂度不同，不直接换算为等量训练价值。	决定数据回流速度

把演示视频当作实验记录，而不是宣传片



字幕式摘要：机器人在真实商业环境中执行接触与搬运相关动作。重点观察：环境是否发生遮挡或扰动、动作是否出现停顿或重试、人工干预如何被记录、最终任务状态如何判定。



实验记录应至少包含

任务：目标物、动作边界、结束条件

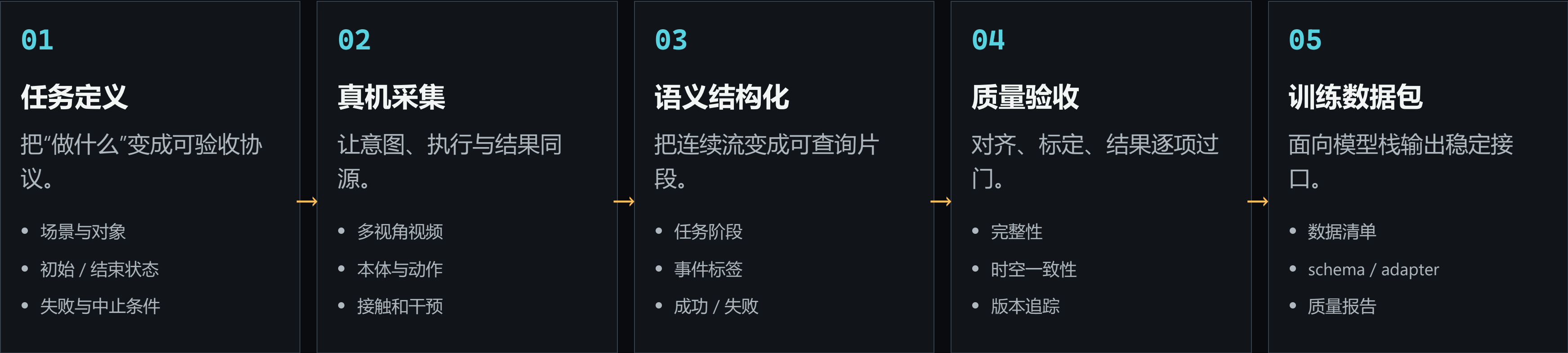
扰动：行人、遮挡、陈列变化、网络波动

失败 / 干预：停顿、重试、接管、恢复片段

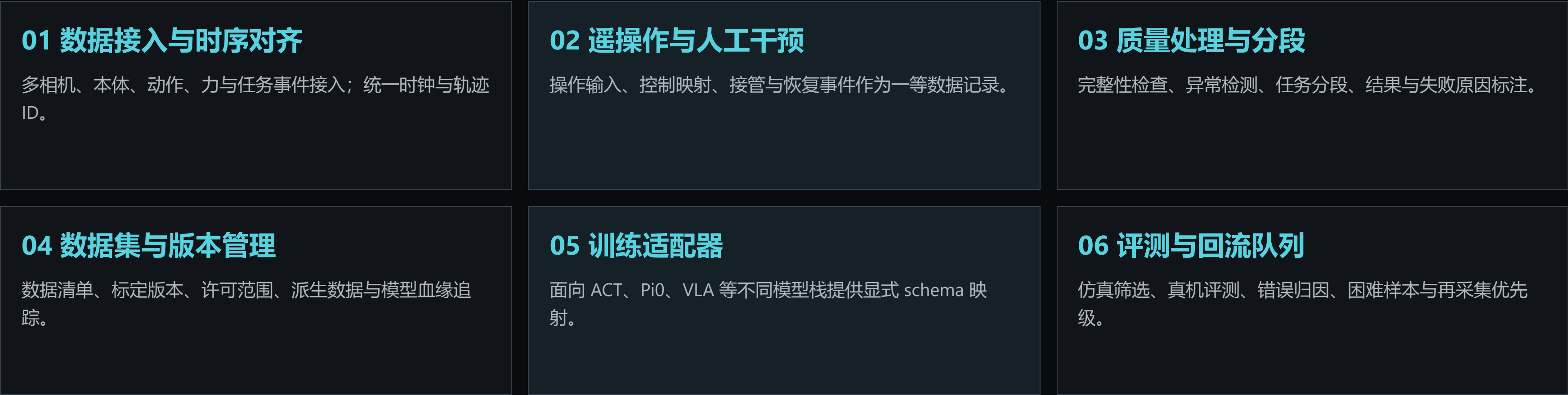
成功标准：物理结果，而非“视频看起来完成”

直接播放：bilibili.com/video/BV1KwQEBxE82

从任务定义，到可交付的数据包



系统边界：采集、治理、训练、评测是一条链



架构价值不在“按钮更少”，而在每次训练结果都能追溯到数据与采集条件。

数据质量不是清洗动作，是进入训练前的门

✓ 信号与文件完整	✓ 时间连续且可对齐
✓ 空间标定版本有效	✓ 动作与状态相互一致
✓ 任务结果可验证	✓ 失败 / 干预 / 恢复可分段
✓ 数据与模型版本可追溯	✓ 发布与使用许可明确

原则： 门槛必须产生结构化拒绝原因，否则“删除坏样本”只会把系统问题藏起来。

REJECTED SAMPLE / QG-TIME-04

拒绝：腕部相机时间戳跳变 83ms



现象： 画面中的抓取接触先于记录动作发生。

风险： 模型会把错误状态当作动作前提。

处置： 从训练集隔离；保留为采集链路回归测试；检查网络缓冲与时钟重同步事件。

训练不是终点，真机评测也要成为数据



真实轨迹 → 训练 → 扩展 → 真机评测 → 归因 → 回流 ↺

训练指标回答

模型在给定数据与损失函数上学到了什么？

评测工件回答

它在独立场景中做成了什么，为什么失败，下一条数据该采什么？

真正难的部分，还没有被数据量抹平

！

跨具身动作表示

不同关节、末端与控制频率如何共享可迁移的动作语义？

！

网络抖动与干预时机

接管事件的真实起点如何在视觉与控制链路上统一标记？

！

数据不平衡与长尾

成功样本很多时，模型是否仍然缺少有解释力的失败经验？

！

隐私与数据治理

部署数据如何完成脱敏、授权、保留期限与派生用途管理？

！

标定漂移

设备移动、碰撞与温漂如何被持续检测，并阻止错误数据入库？

！

长程任务分段

子任务边界、恢复动作和延迟结果怎样形成稳定标注？

！

离线指标失真

模仿损失降低，为什么常常不能预测真实机器人成功率？

！

复现边界

内部测量如何逐步升级为包含设备、负载、样本量的公开基准？

三条结论，以及可以共同定义的接口

01 轨迹是同步的物理事件记录

视频、状态、动作、接触和结果必须共享可追溯的时间与空间基准。

02 质量在采集现场被创造

同步、标定、任务协议和干预记录，无法只靠后期清洗补齐。

03 部署、训练、评测是一条回路

每次真机失败都应该能回到数据队列，成为下一轮采集优先级。

可与社区继续讨论的四类工件

SCHEMA

最小轨迹字段与事件词表

REPORT

同步 / 标定测试报告格式

BENCHMARK

失败、干预与恢复的评测协议

GOVERNANCE

真实部署数据的脱敏与许可清单

本次公开：演示稿、脱敏 schema 片段与质量门示例。未承诺尚未准备好的数据集或代码仓库。

一手来源与内部测量说明

公开研究与项目

[Open X-Embodiment · project & paper](#)

[DROID · project](#)

[DROID · arXiv paper](#)

[Universal Manipulation Interface · project](#)

[Universal Manipulation Interface · arXiv paper](#)

[NVIDIA Isaac GR00T N1](#)

[NVIDIA Cosmos](#)

[Google DeepMind Genie 2](#)

灵御内部实测

时间同步、空间定位、图像传输、本地控制、底层控制、力控范围与小时吞吐，均来自灵御现有 BP3 工程材料。

本稿已注明测试口径及公开限制；这些数字不表示独立第三方认证。后续公开复现时应补充设备、负载、网络条件、样本量和统计分布。

演示稿

[下载 PDF 版本](#)

网页：darkthree.com/speech2/

内容许可：CC BY 4.0（第三方图片、视频与引用材料除外）