

 基元 Prim

公司简介

北京基元智航科技有限公司

组织全球真实世界经验，构建机器人时代的认知基础设施。

不是单一数据、标注、机器人或模型公司，

而是把数据、模型、评测、社区和伙伴连接为可增长网络。

引言

一切智能始于感知。

人类理解世界的方式从来不是靠标签和分类。一个孩子拿起杯子，指尖感受到杯壁的凉意，眼睛看到水面的晃动，大脑在这一瞬间将视觉、触觉、运动信号编织成同一个经验。这个经验不需要被标注，不需要被翻译—它本身就是理解。人类文明的一切技艺，从打磨石器到操控宇宙空间站，都是这样通过第一视角的感知与行动代代相传。

今天，人类站在一个文明量级的跨越点上。从地球文明到星际文明。SpaceX 和 中国航天的可回收星舰、火星殖民计划、太空数据中心、月球基地... 人类的活动半径，正在从一颗行星扩展到一个星系。

而支撑这个跨越的底座，是一个同样宏大的数字。2035 年，全球机器人保有量预计突破 100 亿台——超过人类人口总数。马斯克的判断更为激进：未来机器人与人类的比例将达到 5:1，总量超过 300 亿台。这不再是一个制造业的细分赛道—这是人类文明史上第一次与另一个智能物种共享物理世界。这些机器人将在工厂里组装产品、在医院里辅助手术、在家庭里照料老人、在月球和火星上为人类和机器人铺设第一块地基。

但这两个跨越——从地球到星空，从百万台到百亿台——面临同一个根本性障碍。机器人今天还无法真正理解世界。它们能识别物体，但不能理解场景；能执行指令，但不能判断意图；能在仿真里跑出 99% 的成功率，但一上真机就抓空、捏碎、对着空气操作。差距不在算法，不在算力——在于它们用来学习的数据。互联网上的文本和图片教会了大模型如何对话，但教会不了机器人如何在物理世界中行动。那需要一种完全不同的数据：人类在真实环境中、以第一视角、自然操作所产生的完整行为记录。

这就是基元要做的事。搭建机器人与人类世界之间的那座桥梁。

这座桥的这一端，是人类百万年进化积累下来的感知与行动能力。桥的那一端，是即将遍布地球和宇宙中的百亿台机器人。基元的角色，就是让这座桥变得坚实 、宽阔、可通行—让人类的手艺能够传递给机器，让机器真正理解人类的世界。

2035 年，当 100 亿台机器人在地球上与人类并肩工作、在太空中建设星际设施，回望此刻 —— 希望他们对这个宏伟壮丽物理世界的理解，从基元的桥梁上开始。

Hello, Robot.

Welcome to the Real World.

公司概况

北京基元智航科技有限公司（以下简称"基元"）是一家专注于具身智能大数据平台服务的创新型企业，致力于成为"全球第一具身数据公司"。公司核心研发团队源自大数据科学与脑机智能高精尖创新中心，在大模型微调工具开发、全链路大数据集成与治理等领域具备领先优势。基元立足具身智能产业链核心环节，通过构建世界最强具身智能数据底座，引领物理世界人工智能的规模化落地。目前，公司已与千寻机器人、亦庄机器人、湖北人形机器人创新中心、58 英才网、华为云、京东云、蚂蚁灵波、工业富联、优必选、一目科技等行业头部企业建立合作关系。

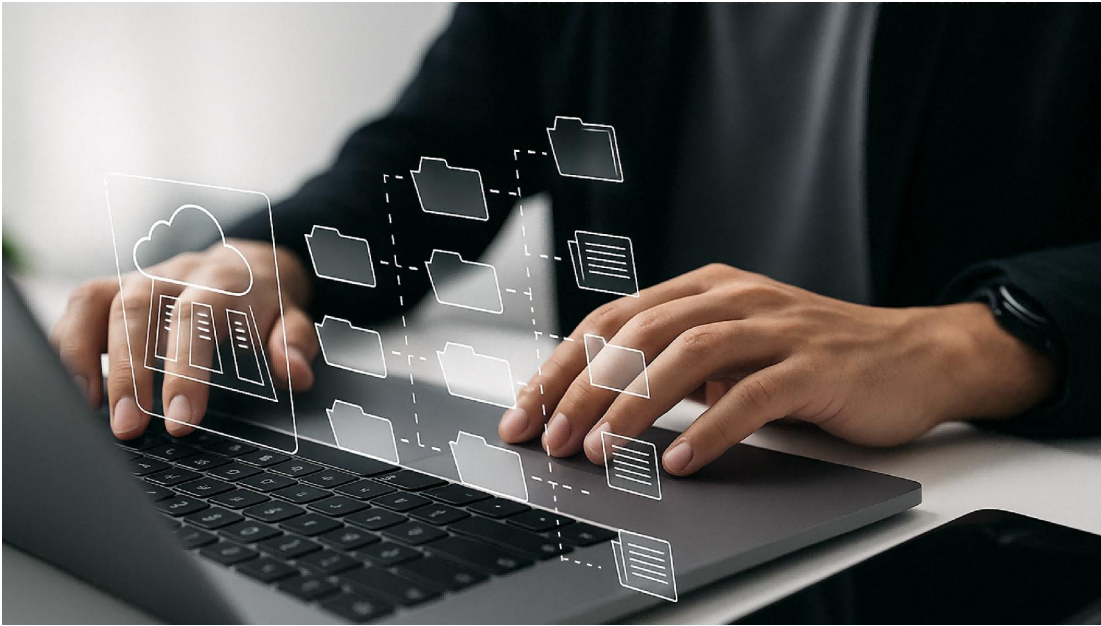
合作伙伴





技术特点

基元的核心技术判断是：数据是与本体、模型并列的具身智能第三极。公司以"无本体人类交互数据"为核心突破口，专注解决机器人技能操作泛化难题，并围绕数据采集、治理、确权、模型训练构建全链路能力。



EGO-人类基础操作 A001

~2万小时人类常见基础操作行为EGO视频（含单目与双目），桌面级操作，包括但不限于“拼牙签、叠放杯子、点蜡烛、打开水杯、将鞋子放进鞋柜”等

[具身智能](#) [视频](#) [EGO](#)

上传者: 柯楠 222MB 6 47 2026/06/24

m_truthfulqa

This dataset is a machine translated version of the TruthfulQA dataset, translated using GPT-3.5-turbo...

[问答](#) [文本能力评估](#) [CC BY-NC 4.0](#)

上传者: sudden3 0B 0 86 2026/05/19

m_arc

This dataset is a machine translated version of the ARC dataset.

[问答](#) [CC BY-NC 4.0](#)

上传者: sudden3 0B 0 21 2026/05/19

m_mmlu

This dataset is a machine translated version of the MMLU dataset.

[问答](#) [CC BY-NC 4.0](#)

上传者: sudden3 0B 0 13 2026/05/19

ManiSkill_PickCube

ManiSkill PickCube Demonstrations Contains the following demonstrations: Motion Planning...

[具身智能](#) [Apache-2.0](#)

上传者: sudden3 0B 0 9 2026/05/19

ISAID航空图像大规模数据集

现有的 Earth Vision 数据集要么适用于语义分割，要么适用于对象检测。ISAID 是第一个用于航空图像实例分割...

[CreativeML Open RAIL-M](#)

上传者: sudden3 0B 0 12 2026/05/19

无感大规模采集 实现高质量超大规模交付

高质量具身数据成本高达1000人民币每小时，数据质量参差不齐，严重制约产业规模化。基元联合产业方，组织百万人以上、真实场景高质量数据，最高日产数据量可达百万小时，从根本上重构了行业的数据供给经济学。

DaaS交付 降低数据使用成本

公司以数据即服务（DaaS）模式向机器人厂商、科研机构及行业应用方交付具身数据，客户无需自建采集体系，按需订阅所需场景、技能类型与数据规模，将具身智能落地的数据获取成本与采集周期大幅降低 90%。

数据确权+RaaS 解决数据交易痛点

采用区块链技术实现数据全链路确权，提供“以结果付费模式”，权益清晰可溯，从根本上解决具身数据质量评价困难、权益归属不清、交易交付成本高这一行业痛点，为数据资产化、高效交易提供基础。

开源社区 构建行业开放生态

公司发起设立“全球具身机器人开源社区”，构建数据集、技能库、模型库及评测体系，以开放生态为全球具身智能开发者提供数据和生态服务。

四位一体生态：数据、模型、评测、开发者的协同飞轮

基元构建的"全球第一"并非单维度的数据规模领先，而是在认知沉淀、标准定义、模型生态和开发者网络四个维度形成全球领导力，构成持续运转的生态飞轮。



数据仓库（Prim Ego Dataset）

目标不是拥有更多数据，而是构建高质量具身数据。提供 Ego 数据、场景数据和高质量具身数据，覆盖机器人每一次操作（取抓放等基础动作）、每一次巡检（工业使用场景）、每一次协作（人机协作、机机协作等），将真实世界的每一次交互都转化为可沉淀、可学习、可复用的经验资产。



模型仓库（Prim Ego Model Hub）

涵盖开源模型、多模态模型及微调/强化模型，支持 VLA 模型、世界模型、动作模型、导航模型及 Agent 模型的开放共享，推动全球具身智能模型生态的协同创新。



评测体系（Prim Ego Bench）

建立针对感知能力、操作能力、协作能力、推理能力、长程任务能力及行业任务能力的行业统一评测标准，回答"哪个机器人更强、哪个模型更强、哪个 Agent 更强"，成为机器人领域的权威Benchmark 生态。



开发者社区（Prim Developer Community）

打造机器人时代的HuggingFace、GitHub 和 ImageNet，汇聚开发者、研究人员、机器人企业、高校实验室及行业专家，形成"数据→开发者→Bench- mark→企业应用→更多数据" 的正向飞轮。

产品体系

从硬件到系统，高性价比丝滑用户体验

高性价比、舒适佩戴的硬件，开放、免费*、功能强大的数据平台

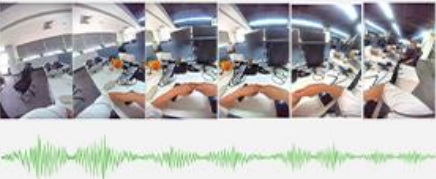
实时预览

提前预览 避免无效采集



时间戳同步

所有传感器
共用一个时钟



部署方式灵活

适配工厂、学校等
规模化部署



第一人称数据采集

Prim_ego Pro



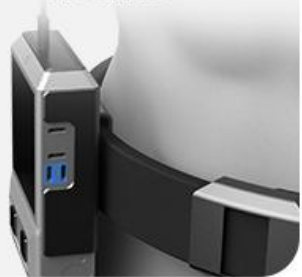
可更换电源

可以多块移动电源
切换使用



扩展友好

支持typeC数据接口



LocalServer
本地部署

端侧清洗
缩小数据体积



六目视觉

真正无感高效采集



佩戴灵活

可腰戴或放于工位旁



无线传输 采传同步

支持边采集边上传



数采设备（Prim Ego Pro / Prim Ego）

在保持最优功能、极致性价比与产品可靠性的基础上，我们深刻尊重和关注每一位数据采集伙伴的使用体验：6个摄像头（Pro版本）可以覆盖横向270度视野范围，大幅度提高有效数据的比例，珍惜数采伙伴的劳动价值，市场最轻量化的结构设计幅度减轻佩戴者肩颈压力，分体式电池避免贴肤发热以及极低概率下电池自燃给佩戴者带来的灼伤风险。

同时，我们提供端侧数据清洗能力，在不影响模型训练结果的前提下，减小数据规模，以及对应模型训练时的算力消耗，同时去除商业敏感与个人隐私信息，保障数据合规。

基元智航
PRIM

技术广场

数据集

任务大厅

模型库(在建)

工具中心

机器人

伙伴专区

设备专区

数据集

技术领域 应用场景 行业 语言 协议

专题

训练阶段

数据类型

文本处理

图像

音频

搜索数据集关键字

共 1070 个结果

EGO-工业手工生产

概述：· 采集场景：工业环境下采集手工任务，100%真实工业环境和工序操作过程；· 任务内容：工业工序下的手部操作动作，涵盖缝纫，拷边，…

视频

上传者：何桢 151MB 0 11 2026/07/17

EGO-居家环游视频

概述：· 采集场景：居家环境下进行真实活动环游任务，100% 真实家庭环境；· 任务内容：真实家庭环境下进行漫游，涉及到客厅、餐厅、厨房等…

视频

上传者：何桢 588MB 0 16 2026/07/17

EGO-酒店-双目+双手IMU

概述：· 采集场景：100% 真实酒店环境；· 任务内容：任务包含房间布草、床务整理、房间清洁等；· 采集规格：双目视觉 2048 * 768, 30fps；…

具身智能 视频 IMU EGO

上传者：何桢 78.6MB 0 4 2026/07/16

桌面系列任务001

0

数据平台（Prim Studio）

基元提供开放、免费*的数据平台，涵盖数据存储、清洗、标注、模型训推、本体部署在内的全链路工具链，Prim Studio与基元的开发者社区密切链接，期待与包括您在内的全球具身开发者共享、共建。

* Prim Studio亦可提供本地部署的商业版本，商业化收益将大部分用于开源社区建设。

解决方案

从数据到智能，从数据到收益，全链路贯通



大规模数据采集与交付

面向全球具身智能企业提供最高每日超万小时的高质量数据集交付，包含 全场景搭建、数据采集（可指定数采设备）、数据存储、数据清洗与标注（可使用指定工具链）服务，可选择定向、定期共享数据以获得服务费折扣。

最后更新：

开始时间

-

结束时间

搜索：

请输入数据集名称

数据集名称	标签	最后更新	数据集权限	是否外部	版本信息	创建者
EGO-工业手工生产	 视频	2026/07/17 13:56:57	数据集权限：公开 允许下载：是 允许讨论：否	否	-	何桢
EGO-家务-A002	 具身智能  视频 ...	2026/07/17 13:12:50	数据集权限：公开 允许下载：是 允许讨论：否	否	数据文件样例：公开，免费 下载	何桢
EGO-居家环游视频	 视频	2026/07/17 11:50:49	数据集权限：公开 允许下载：是 允许讨论：否	否	采集样例：公开，免费 下载	何桢
EGO-工业-单目&双目	 具身智能  视频 ...	2026/07/16 23:37:32	数据集权限：公开 允许下载：否 允许讨论：否	否	-	何桢
EGO-酒店-双目+双手IMU	 具身智能  视频 ...	2026/07/16 22:45:48	数据集权限：公开 允许下载：是 允许讨论：否	否	数据样例：公开，免费 下载	何桢

数据资产开发与数据托管服务

面向各类产业企业、事业单位，提供专有具身智能数据集建设、企业垂类模型开发、数据托管、数据资产运营销售服务。面向数据和人力服务公司，提供数采硬件租售、专用数据平台建设、数据托管、数据运营销售服务。



具身智能开发者套件 Prim Devkit

面向科研机构、个人开发者、教育院校用户，包含可以便捷使用的数据集、数据工具链、模型微调工具链、机器人、数采设备和必要配件的具身智能开发套件。支持用户开发者套件基础上进行科学研究、教育教学以及主流Benchmark评测。



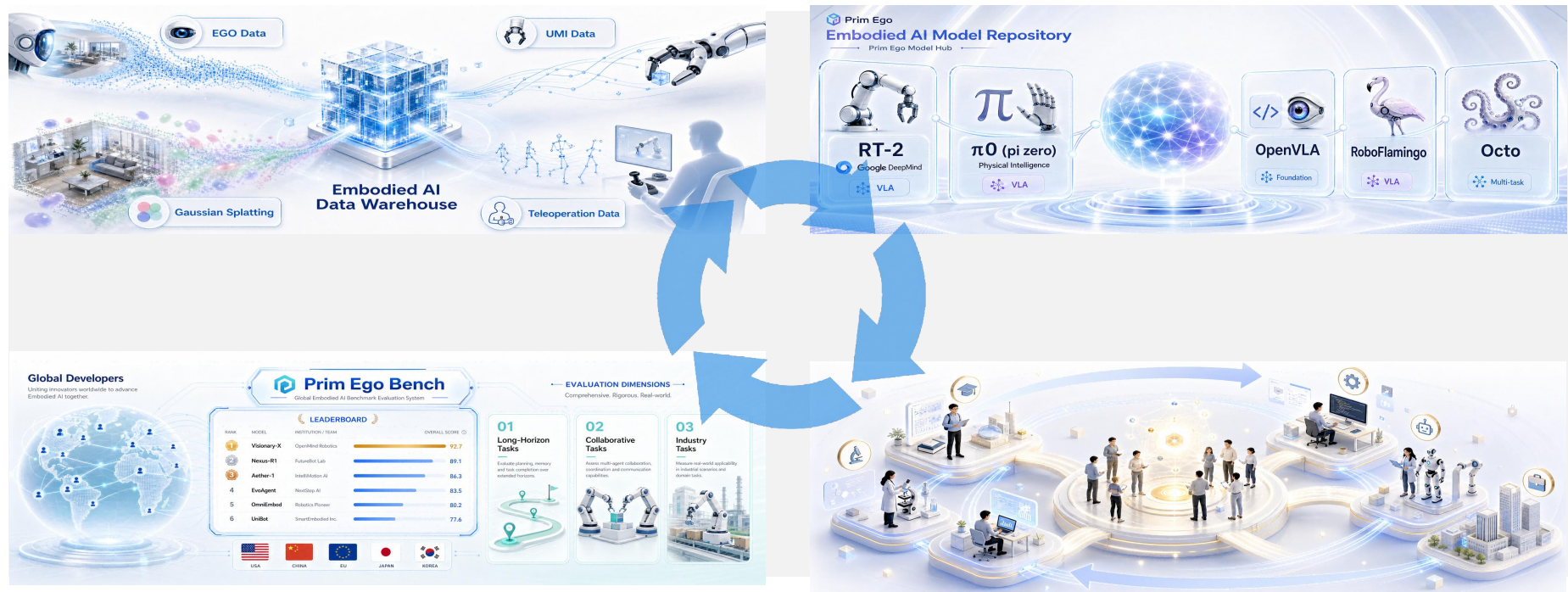
1150数据工厂

面向各类数采基地业主单位，提供 1 真机数据工厂 + 1 数据和算力平台 + 50+分布式产业企业（1150）数据工厂建设与运营服务，协助业主单位实现具身数据资产运营销售与确权、数算一体Token销售。



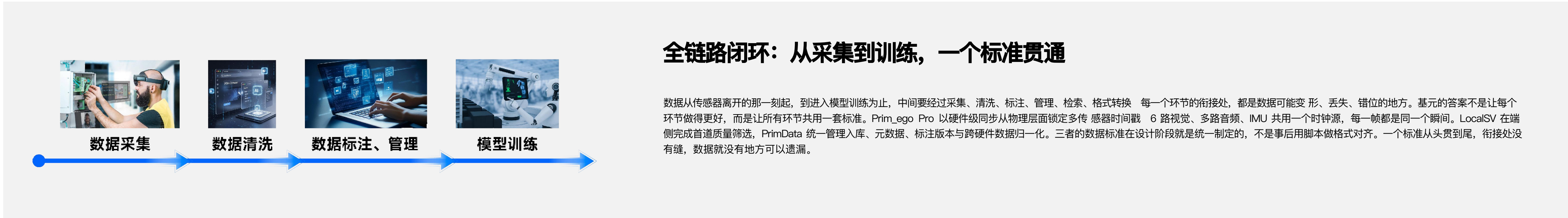
核心优势

什么让我们不一样



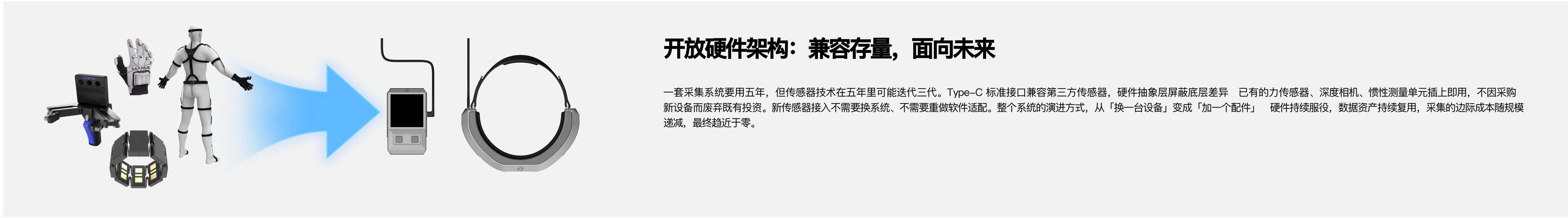
网络效应：四位一体生态

具身智能的数据、模型、评测、社区，本质上是同一个问题 让机器人理解物理世界。把它们拆开交给四个厂商，每个厂商只优化自己那一环，最终拼接出来的是一条处处漏水的管线。基元的EGO基础设施，把四者装进同一个飞轮：Data Engine 产出的标准化数据直接喂给模型仓库，Benchmark的评测结果反哺数据采集方向，开发者社区让每一环的参与者都能为整个系统贡献和受益。飞轮一旦转起来 数据越多，模型越好；模型越好，评测越准；评测越准，数据方向越清晰；方向越清晰，参与者越多；参与者越多，数据越多 每一圈都在拉大优势。这不是堆功能，是让四个齿轮互相咬合，产生单点方案无法复制的复合增长。



全链路闭环：从采集到训练，一个标准贯通

数据从传感器离开的那一刻起，到进入模型训练为止，中间要经过采集、清洗、标注、管理、检索、格式转换 每一个环节的衔接处，都是数据可能变 形、丢失、错位的地方。基元的答案不是让每个环节做得更好，而是让所有环节共用一套标准。Prim_ego Pro 以硬件级同步从物理层面锁定多传 感器时间戳 6 路视觉、多路音频、IMU 共用一个时钟源，每一帧都是同一个瞬间。LocalSV 在端侧完成首道质量筛选，PrimData 统一管理入库、元数据、标注版本与跨硬件数据归一化。三者的数据标准在设计阶段就是统一制定的，不是事后用脚本做格式对齐。一个标准从头贯到尾，衔接处没有缝，数据就没有地方可以遗漏。



开放硬件架构：兼容存量，面向未来

一套采集系统要用五年，但传感器技术在五年里可能迭代三代。Type-C 标准接口兼容第三方传感器，硬件抽象层屏蔽底层差异 已有的力传感器、深度相机、惯性测量单元插上即用，不因采购新设备而废弃既有投资。新传感器接入不需要换系统、不需要重做软件适配。整个系统的演进方式，从「换一台设备」变成「加一个配件」 硬件持续服役，数据资产持续复用，采集的边际成本随规模递减，最终趋近于零。

Thank you
for watching.

让机器人真正像人 一样理解世界

Hello, Robot. Welcome to the Real World.