

智能座舱AI Agent 三大演进趋势及案例分析

2026年5月

北京佐思信息咨询有限责任公司



更多免费行业报告

并购家

www.ipoipo.cn

目录

1

智能座舱AI Agent演进趋势

2

3大案例分析

演进趋势一：智能座舱情感引擎化，从“执行任务”的机器系统向有温度、懂陪伴的数字伙伴演进

- 智能座舱“情感引擎化”是指，智舱通过多模态感知技术，实时识别并理解驾乘人员的情绪状态（如喜悦、疲惫、焦虑、分心），并基于此主动提供情感适配的交互内容、环境调节和关怀服务，让车辆从被动执行指令的“机器工具”，升级为能主动感知、理解、共情并响应驾乘者情绪与身心状态的“有温度的数字伙伴”。
- 智能座舱向情感引擎化演进，本质是赋予机器“情商”，让它不再是冷冰冰的工具，而是一个试图理解、适应、关心用户的出行伙伴，这反映了用户对汽车作为“移动生活空间”情感价值的深层需求。

智能座舱向情感引擎化演进的驱动力：

- 1

- 在硬件参数同质化的时代，“情感”成为最高级的差异化。一个“温暖贴心”的座舱人格，能形成强烈的品牌认同和用户粘性，让车从“配置单”变成有“性格”的伙伴。这种体验一旦形成，品牌忠诚度会远高于配置带来的好感
- 2

- 情感引擎是主动安全的高级形态。人在驾乘场景下，会出现大量的情绪场景，如早高峰烦躁路怒、加班晚归疲惫低落、长途驾驶犯困无聊以及带娃出行吵闹焦虑等。
 - 通过识别出驾驶员的愤怒、分心或多种情绪，及时干预（如深呼吸引导、播放平静音乐），能预防事故，智能座舱补上了座舱缺失的“情绪价值”，提升了旅途的安全性、舒适度，给用户带来愉悦感
- 3

- 实现车 → 家 → 办公等跨场景体验及跨终端的无缝流转最顺滑的连接方式，就是情感与习惯。让情感成为一条贯穿全天的数据流，而不是一个个孤立指令；持续的情感交互产生了极具价值的“用户情感数据”，在合法合规前提下，这能用于优化产品，甚至未来可能衍生出个性化的情感关怀订阅服务、健康监测服务等



打造差异化
品牌人格

情感需求是
刚需

情感是跨场景、跨
终端流转的最好的
纽带

01 情感感知输入

多模态情绪感知

- **视觉**：车内摄像头（通常位于A柱或方向盘后）通过微表情识别（嘴角弧度、眉毛动作）、面部血谱分析、视线追踪、姿态识别（是否频繁揉眼、低头）来判断情绪和注意力。
- **听觉**：麦克风阵列分析语音的语调、语速、音量，甚至非语音声音（叹息、哈欠），进行情感计算。
- **生理信号**：通过方向盘或座椅内置传感器，监测心率、皮肤电反应等，客观反映压力或兴奋水平。
- **长时记忆与周围环境理解**：通过记忆网络、上下文窗口管理能力记住用户常去地点、音乐喜好、情绪触发点（如讨厌拥堵），并结合时间（深夜长途）、地点（拥堵路段）、车辆状态（激烈驾驶）等信息，综合判断情绪倾向，避免机械式回复

02 情感理解与计算

情感计算中枢

- **跨模态融合算法**：视觉→面部嵌入；音频→语音情感嵌入；文本→语义嵌入；融合输出情绪状态（主情绪 + 强度 + 置信度）；
- 以广汽2026发布的ADiGO Intelligence端云一体架构多模态情感表为例，在获得多模态情绪感知后，将基于输入数据，输出标准化情绪标签（开心、哭闹、愤怒、激动、中性等），即进行情绪分类输出，为后续表达提供明确情感导向，核心是“精准识别，不误解情绪”。

广汽2026发布ADiGO Intelligence端云一体架构多模态情感表



03 情感表达与决策输出

拟人化情感表达与执行

- **语音交互**：这是核心。情感化TTS（语音合成）能让车机的声音充满关切、欢快或沉稳。对话逻辑不再是Q&A，而是具有共情能力的多轮交流。
- **多通道氛围联动**：情感引擎会驱动座舱的“五官”协同表达。例如，在庆祝场景下，联动播放激昂音乐、调亮氛围灯、释放愉悦香氛；在需要专注时，自动调暗屏幕、关闭娱乐音源、收紧安全带。
- **数字/实体形象**：如商汤绝影推出的可悠全场景多设备协同智能终端、蔚来的NOMI，通过实体生动的表情，将情感交互**视觉化、人格化**，极大增强情感连接的亲切感。



蔚来Nomi、小米超级小爱及商汤绝影可悠智能座舱情感引擎化的技术布局对标

对标分析	蔚来 NOMI	小米汽车 超级小爱	商汤绝影 可悠 (Care U)
核心定位	车载专属情感伙伴 (懂我、陪伴)	人车家全生态 AI 助手 (全能、好用)	全场景多设备协同智能终端 (关心、懂家)
技术底座	全栈自研 NOMI 情感引擎 + NOMI GPT	MiMo大模型 + 澎湃 OS	自研空间多模态交互系统 + 全场景多设备智能体协同体系
拟人化 · 音色 / 语调	- 支持MBTI人格自定义 - 拟人化·音色/语调 ✓ 支持温柔/冷静/活泼等多种人格 ✓ 语调自然停顿、抑扬顿挫 ✓ 高拟人度	升级版蜜糖音色 (2026 最新) - 强拟人听感：自带思考停顿、语气起伏、情绪温度 - 特色人格模仿：林黛玉、傲娇、活泼等多种腔调 - 多方言：粤语、四川话、东北话等自由切换	多音色可选 + 场景化语调 - 基础情感音色 (温柔 / 沉稳 / 活泼) - 语调随场景 / 情绪适配 (关怀 / 提醒 / 闲聊) - 清晰自然, 强调 “安心、亲切” 的家庭感
拟人化 · 视觉 / 动作	强具身实体 (NOMI Mate 3.0) 340° 精细旋转 (400 步静音电机) 100 + 动态表情 (AMOLED 高清屏) - 音乐律动：随 15 种曲风节奏摇摆 - 情绪联动：夸时脸红、逗时委屈、认真时专注	虚拟萌宠 + 氛围联动 - 车机虚拟萌宠形象 (德文猫、水獭等) - 表情 / 动作反馈、语音播报时动态嘴型 - 氛围灯、音效、屏幕动效协同	独立实体具身 360° 旋转云台 + 震动反馈 50 + 灵动 OLED 表情 (简洁可爱风) - 姿态、灯光、声音三维同步反馈
记忆体系	座舱级双层记忆 - 短期上下文 - 长期个人偏好 (音乐 / 温度 / 日程) NIO蔚来汽车	生态级跨端记忆 - 手机 / 车 / 家数据互通 - 长会话上下文记忆	生长型家庭三级记忆 - 感知记忆 → 事实记忆 → 认知记忆 - 全家成员统一画像、全场景同步
情感感知	语音语调 + 人脸微表情 + 驾驶状态 (疲劳 / 烦躁) NIO蔚来汽车	语音 + 人脸 + 方向盘生物信号 (高阶)	语音 + 人脸 + 手势 + 空间姿态 + 环境
主动关怀	疲劳安抚、节日祝福、个性化问候NIO蔚来汽车	日程提醒、天气关怀、生态设备联动	家庭场景关怀 (老人 / 儿童)、跨场景提醒
场景边界	车内深度闭环 (车机 → NIO Phone → 手表) NIO蔚来汽车	人车家全域 (车 → 手机 → 米家 → 平板)	车 → 家 → 办公全链路 (车载 New Member → 家用可悠)
拟人化核心优势	情感沉浸最强 实体机器人 + 专属声线 + 丰富表情动作, 最像 “副驾伙伴”	人格化与趣味最强 蜜糖音色 + 角色模仿 + 方言 + 萌宠, 交互最生动好玩	家庭共情最强 跨端统一记忆 + 全家认知 + 温和反馈, 像 “家庭新成员”
载体形态	车载实体 + 虚拟	车机虚拟 + 手机小爱 + 生态联动	独立硬件实体 (可悠) + 车载 / 桌面双场景

生长型家庭认知记忆

- 三层记忆体系：
 - a) 感知记忆：记录对话、行为、环境细节（瞬时 / 场景记忆）。
 - b) 事实记忆：提取偏好、日程、关系图谱（长期记忆）。
 - c) 认知记忆：理解家庭互动模式与潜在需求，动态更新用户画像。最后实现跨设备统一，即上下文、任务进度、偏好实时同步，杜绝换设备“断片”。

精准感知与身份绑定

- 多维度识别人、关系与偏好：50 声纹 + 人脸区分家庭成员；通过交互学习建立个人偏好（音乐、内容、作息）；支持 50 + 表情与震动反馈，传递情绪状态。



情感适配与全场景协同

- 主动情感适配：疲劳 / 焦虑时调音量、播舒缓内容；老人久坐提醒、孩子上课提醒。
- 全场景接力：家庭→车载→办公无缝续接，如家中定路线上车即导航，加班纪要同步电脑。
- 技能自进化：支持技能动态拓展（无需 OTA），强化学习持续优化决策逻辑

情感引擎的技术底座



- 感知层：捕捉语音、人脸、手势、空间姿态
 - a) 语音：双麦阵列（180° 拾音、50 声纹 ID）
 - b) 视觉：200 万像素摄像头
 - c) 姿态：100Hz 六轴陀螺仪
- 端云协同架构
- 多模态大模型（日日新 SenseNova V6）、类人记忆框架、端侧原生智能体框架（支持 MCP），为情感与认知提供底层支撑
- 反馈层：1.75 英寸可 360° 旋转 OLED 屏（50 + 表情）、震动与灯光联动，实现拟人化情感表达
- 以“你”为中心实现对家庭全成员的精准记忆与深度理解，形成用户和家人的画像，记住用户及其家人的习惯、偏好、禁忌、雷点

- 车载AI智能体的情感引擎的核心之一在于**生长型记忆体系**，它是AI智能体了解用户、读懂用户真实需求的基础设施”，人类情感的精准判断不仅仅依赖单次交互，更取决于长期记忆的多次累积。
- 如果车载 AI智能体记不住，它就只能每次用户说啥，它做啥，但每次都像第一次认识用户，无法共情、无法预判、无法体贴。
- 商汤绝影AI智能体可悠的三层记忆体系，可以实现感知记忆，如记最近 30 分钟 / 本次行程发生了什么，让对话从“机械式”→“温情主动式”；
- 可悠的另一大差异化优势在于**家庭认知记忆，打破当前主流 AI 记忆体系仅适配单人工作场景**、仅做流水式信息存储的行业局限，打造了一套 AI 驱动、会自主生长的认知记忆引擎，通过独有的专属记忆体系，以“你”为中心实现对家庭全成员的精准记忆与深度理解，形成用户和家人的画像，记住用户及其家人的习惯、偏好、禁忌、雷点，产生一种被注意、被关心的体验；

趋势二：行业加速打通出行-生活-办公全场景，全场景跨终端的系统级AI智能体将成为下一个风口

➢ 车不再是终点，而是进入全场景 AI 的入口与移动中枢。智能汽车的下一站，不仅仅比谁的车更智能，而且比谁的 AI 能更快、更好的覆盖用户从车到家、到办公、到生活的每一分钟，并且以具身智能（实体机器人 / 数字分身）的形态，成为真正陪伴、记忆、主动办事的“生活伙伴”

用户需求的再演进：24小时*7天的随身AI智能体伙伴



汽车是全场景跨终端AI智能体的其中一个服务节点

- 用户对汽车的期待，已超越A点到B点的位移。它被期望是办公室、客厅、休息室、娱乐中心的延伸。
- 一个只能在车内好用的AI，无法满足用户在生活中无缝切换的需求。用户需要的是一个“随身的智能伙伴”，车只是其一个重要的服务节点



竞争升维：从“产品功能战”到“生态体验战”

- 竞争焦点将从“我的车有什么功能”转向“我的车能为我的全生活提供什么价值”
- 未来当单一车辆的体验触达天花板（三电系统、智驾标配化、智舱AI模型趋同化后），必须将价值延伸到用户24小时的生活场景中，构建更深的用户粘性和品牌护城



从“人找服务”→“服务随人”

- 当前智能座舱与移动终端之间的数据壁垒尚未打破，导致用户体验呈现显著的“记忆割裂”与“场景断层”。车机局限于行驶数据而无法延续用户在家庭与办公场景中的上下文记忆，手机则难以同步座舱内的行程动态与偏好变更。由此，人工智能沦为多个独立运行的“工具集合”，用户不得不在切换设备时反复进行指令重置与信息复述，交互体验呈现出碎片化、断点式的割裂状态。
- 未来的智能交互应实现从“人找服务”到“服务随人”的范式跃迁。依托统一的跨端情感计算与数据流，用户的无感身份与情境记忆将在手机、座舱、家居与办公空间之间无缝流转。AI将打破单点设备的物理边界，以连续、主动、无断点的服务姿态，构建起一个始终在线、懂你所需的全场景伴随式智能体。



从商业价值来看，打通“车-家-办公-个人”大空间，构建更大更广的商业闭环

- 把「车」的能力延伸到「家」和「个人以及办公」，形成「车-家-人-办公」全场景覆盖，从「车载供应商」升级为「全场景 AI 服务商」，打开了 To C 的增量市场

三类以车为入口，加速布局全场景、跨终端的系统级AI智能体的参与者

三类以车为入口，加速布局全场景、跨终端的系统级AI智能体的参与者

第一类：生态构建者

“生态构建者”将提供更流畅的体验

- 代表玩家：小米、华为
- 特点：拥有或深度掌控硬件、操作系统和关键应用生态，旨在打造一个体验闭环的智能体
- 以小米为例
 - ① 核心产品/架构：手机、汽车及各类智能终端产品、AI大模型基座、适配各类终端的端侧AI大模型、澎湃OS（跨端底层系统）、“人车家全生态”战略。
 - ② 布局特点：以手机为个人中枢，以汽车为移动空间，以智能家居为生活场景，通过自研的MiMo大模型和澎湃OS，实现数据和服务的无缝流转。miclaw是其实验性的核心智能体原型

第二类：技术赋能者

加速全行业的智能化普及

- 特点：以AI技术和开放平台为核心，为车企和硬件厂商提供AI智能体能力，旨在构建一个广泛的联盟生态；
- 代表玩家：商汤绝影
 - ① 核心产品/架构：“日日新”大模型体系、“可悠”全场景多设备协同智能终端、OpenClaw（智能体开放框架）。
 - ② 布局特点：依托强大的CV和多模态大模型技术，推出“可悠”作为面向家庭、出行、办公场景的智能体。通过OpenClaw协议，致力于连接不同品牌的汽车、家电、机器人等设备，构建开放的全场景AI网络。
 - ③ 优势：AI技术领先，姿态开放，适合与众多希望“快速智能化

第三类：车企自研派

- 以理想、蔚来等为代表的车企自研派
- 这些车企均在自研大模型（如理想Mind GPT）和座舱智能体，并积极与手机、家居生态联动。它们的目标是先成为自己生态内的“垂直整合者”，再考虑开放。

蔚来Nomi、小米超级小爱及商汤绝影可悠全场景AI智能体布局对标分析

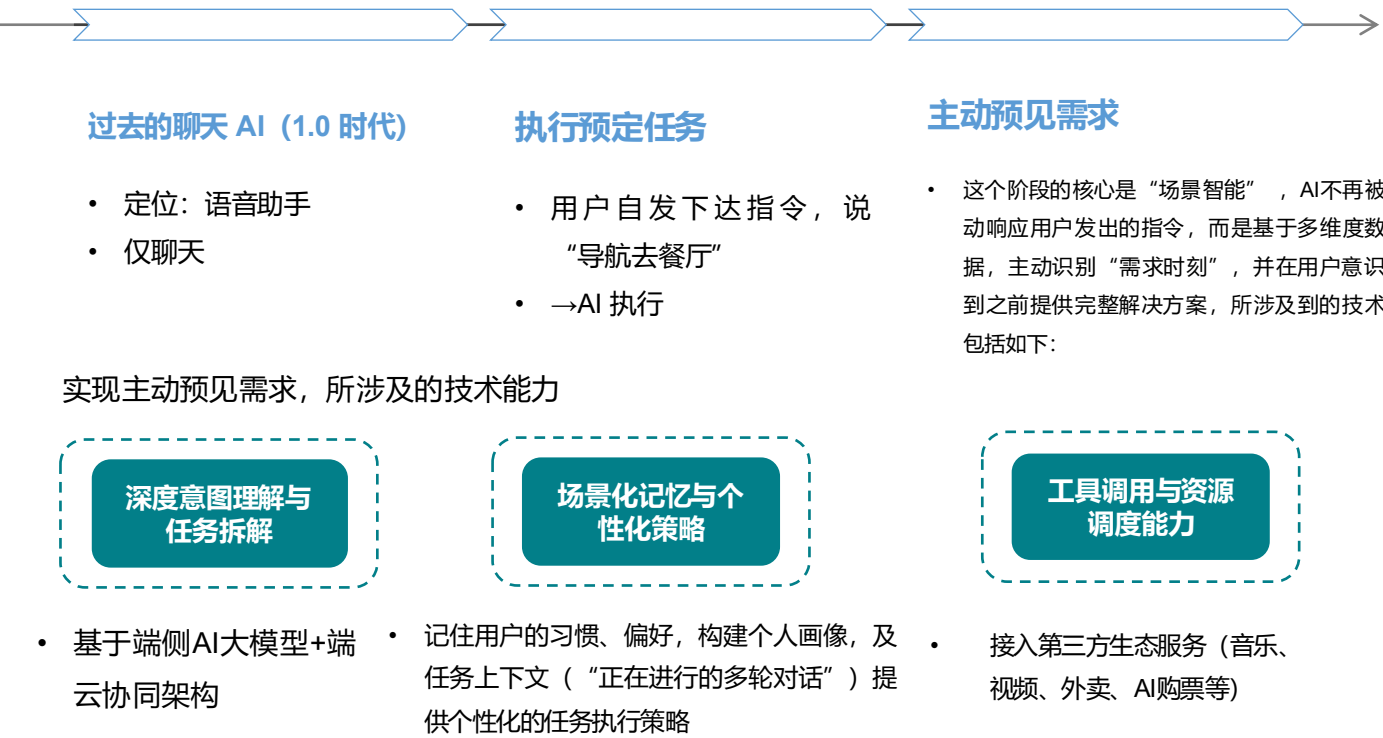
维度	商汤绝影	小米	蔚来汽车
核心定位	技术供应商与连接器	生态整合商与制造商	高端服务商与社区运营者
战略路径	开放协议，联盟赋能	垂直整合，闭环体验	车为核心，服务延伸
核心载体	“可悠”（桌面具身机器人）	“超级小爱”+澎湃OS+全生态硬件	NOMI
记忆体系	生长型三层记忆（感知/事实/认知），全局统一同步	跨会话长记忆+全生态用户画像，OS统一调度	长时空记忆+车规级隐私安全防护
跨端协同	LinkCore 协同中枢，打通车/家/办公/个人设备	HyperConnect 跨端框架，手机/车/家/ IoT 无缝接力	神玑芯片+ SkyOS 全域调度，车机/ NIO Phone / 家居联动
AI 底座	商汤绝影New Member座舱同源AI模型	MiMo 模型系列	NOMI GPT
技术架构	端云协同	端云协同	端云协同
落地场景	车→家→办公全链路，强调“同一个灵魂，不同分身”	车/家/移动/办公，以地理围栏触发主动服务	以驾驶为核心，延伸至车内服务、车家互联、出行服务微博
优势场景	快速为多品牌车企/家电赋能	跨设备无缝协同、主动智能	车联服务、情感化陪伴、圈层专属
主要挑战	体验深度、生态碎片化	生态封闭性、持续投入压力	生态广度、技术通用性、成本控制
差异化优势	以实体机器人突破座舱边界。依托座舱 AI 技术积累，通过可悠硬件将服务延伸至家庭与办公，主打“具身智能+跨设备协同”，为车企提供差异化全场景方案，合作模式以“技术输出+生态共建”为主	以 OS+统一AI模型底座+全生态终端布局，通过超级小爱串联手机、汽车、智能家居，打通全场景	NOMI Agent 侧重车内与车家互联的服务闭环，技术壁垒集中在车规级稳定性与智驾决策能力

趋势三：智能座舱任务化深度解析，从聊天 AI 到靠谱执行者的范式跃迁

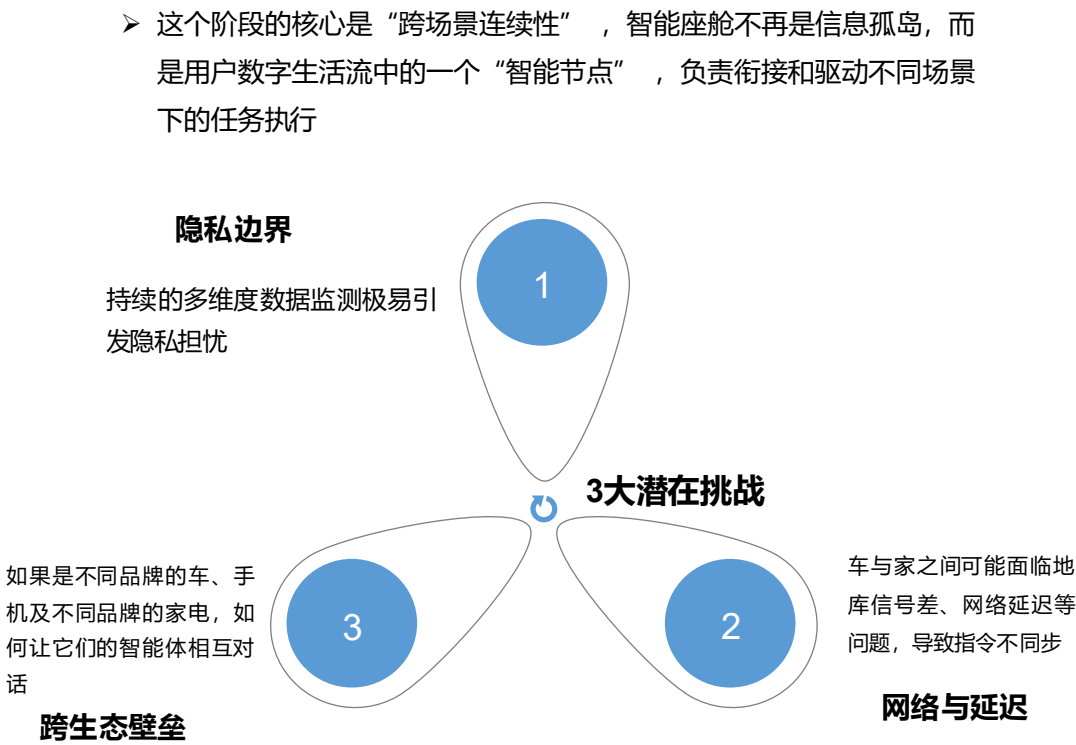
- 智能座舱“任务化”是指，智能座舱的AI系统从一个被动响应指令的“聊天机器人”，转变为一个能主动理解用户复杂意图、自主规划分解步骤、并调用一切可用资源（车内功能、云端服务、生态设备）来可靠完成任务的“智能Agent”。
- 座舱AI智能体，如果只停留在“能聊天、能点歌、能开空调”的阶段，同质化严重。用户真正的痛点，不仅仅是“和车说话”，而是“用车解决问题”，比如通勤路上整理纪要、出行前规划行程、回家前联动家里设备等等。从“满足交互爽点”，转向“解决用户痛点”

从形态来看，任务化演进的趋势将朝着2个方向协同发展

- 方向一：从“执行预定任务”到“主动预见需求”（即从“人找服务”→“服务找人”，实现一句话完成任务闭环



- 方向二：从“车内任务”到“生活流任务”



案例分析1：可悠接入OpenClaw，升级为“能动手、会规划、跨设备、全场景执行的 AI 智能体”

➤ 可悠接入OpenClaw，本质上就是为这个贴心的伙伴装上了一对能高效处理任务的“龙虾钳”，让它从一个善解人意的AI 陪伴者，蜕变为一个能帮你主动完成实际工作的“家庭新成员”。



案例分析2：小米汽车智能座舱Agent小爱定位及关键能力的演进

懂情绪，能思考

小米汽车智能座舱Agent小爱定位及关键能力的演进

车载语音助手：能听懂

高效交互合作伙伴：免唤醒，更主动

超级小爱 2.0

2024.03

初代座舱小爱（SU7 首发）

- 定位：车载语音助手
- 端侧模型：搭载MiLM-1.3B车载大模型；
- 关键能力：
 - ① 五音区精准识别：区分座位权限（主驾控行车功能，乘客控空调 / 车窗）。
 - ② 基础车控 + 车家互联：覆盖 90% 座舱功能，手机 PIN 到车机、联动米家设备。
 - ③ 车外基础唤醒：支持车外简单指令，防御陌生人误触。

2025.09

HyperOS 1.10.0

- 指令 / 昵称免唤醒：指令免唤醒（如“导航去公司”）、昵称免唤醒（指令前后加“小爱”）；
- 多模态进阶：新增“问建筑”功能，即可识别车前建筑颜色 / 高度，支持圈选继续识别；
- 视觉与音色焕新：天际屏语音编辑，新增“柚子”等播报音色

2025.12

HyperOS 1.11

- 定位：主动智能“贴心管家”，根据用车行为与环境主动问候、提建议（如恶劣天气提醒湿滑模式）。
- 语音任务创建：一句话生成提醒 / 任务（如“后排没人时关空调”）。
- 车外功能扩展：支持露营模式、哨兵模式、车外拍照等。

2026.04

- 定位：主动智能体，以XLA 认知大模型为核心，实现“拟人化 + 全场景 + 人车家贯通”
- 搭载XLA 认知大模型；MiMo-V2 系列大模型
- 关键能力：
 - ① 拟人交互：情绪感知（疲惫时主动安慰）、跨会话长记忆（记住偏好路线）、方言 / 角色腔调（上海话、林黛玉语气）。
 - ② 全场景导航：模糊语义搜索（“附近最大滑雪场”）、多途经点串联、动态路线调整。
 - ③ 车外交互成熟：360° 拾音、声纹 + 数字钥匙双重验证、车外语音泊出（窄车位神器）。
 - ④ 人车家深度贯通：车机 / 手机 / 米家统一账号，在家控车、在车上看家中摄像头

目录

1

智能座舱AI Agent演进趋势

2

3大案例分析



商汤绝影：可悠AI智能体分析

可悠基本介绍	
定位	打通 车、家、办公 三大空间的统一智能体、全场景跨设备协同智能终端、家庭AI新成员
使命	认识你、了解你、关心你
形态	小巧桌面机器人，柔软亲肤外壳，1.75 英寸 OLED 屏，可360° 旋转，能显示 50 + 种灵动表情
核心技术	源自商汤绝影车载 New Member 座舱 AI 系统，是其能力的实体化、全场景延伸 - 生长型家庭认知记忆框架：三层记忆（感知→事实→认知），能识别家庭成员、理解关系、学习习惯，真正“越用越懂你” - 全场景智能体协同：场景无缝衔接+与车载 New Member、手机、电脑共享记忆与任务 - 空间多模态交互：支持语音、视觉、姿态、震动多维交互 - 端云协同架构：端侧快响应，云端强推理，兼顾速度、隐私、智能
三大应用场景	车载场景：出行规划与车家互联
	办公场景（接入 Open Claw 工具链）：类比桌面“贾维斯”
	家庭场景：家庭助手与健康关怀



➤ 可悠是商汤绝影于2026年4月8日推出的、面向车 — 家 — 办公全场景的具身智能机器人，以统一记忆与端云协同为核心，是具备感知、认知、表达、行动能力的生长型家庭智能体



绝影打造可悠的背后推动力因素

突破车载 AI 的“场景天花板”

1

- 用户一天在车里只有1~2小时，剩下22小时都在家、办公场景等。如果只做智能出行场景，抓住用户的碎片时间非常有限，无法形成持续交互数据，大模型的能力无法充分释放；
- 此外，整车品牌可以通过可悠将其在智能智舱验证过的AI能力，移植到家庭场景-办公场景

从“车内”延伸到“全场景”
车载AI智能体→全场景AI智能体

以人为核心，打造统一AI智能体，
打破割裂体验、形成跨设备流转

2

- 目前，行业面临车机 AI、家居 AI、手机 AI、办公 AI 互相独立，记忆不通、身份不通、服务不通。车里设的提醒，家里不知道家里的习惯偏好，车机不继承换一个设备就要重新说一遍需求；
- 可悠的战略意义，就是把“人”作为中心，基于统一的AI智能体，实现用户统一身份、统一记忆、统一意图理解和统一服务输出

AI 范式的升级：以人为中心的AI
智能体

抢占“车+家+办公”多核心入口，构建下一代智能生态

3

- 智能生活的三大核心高价值场景为家庭、出行、办公。家庭是生活与情感中枢、IoT 控制入口；汽车是“第二客厅”与高价值流量入口；办公场景承接用户每日 8 小时核心时长，是车家智能体验向外延伸的关键补充。
- 目前行业各家的路径如下：
小米：IoT 为主，车是延伸
华为：鸿蒙系统为主，多端流转
其他车企：车机为主，家是辅助

绝影以“AI 智能体”为核心，统一车+家+办公



可悠核心技术1：空间多模态交互技术解析（1）

可悠能在复杂环境中精准识人、感知环境，甚至预判用户意图，核心依托商汤绝影自研的**空间多模态交互系统**。这套系统覆盖“**精准感知—深度理解—灵动交互**”三大核心环节，构建了从信息捕捉、智能决策到互动反馈的全链路闭环，是可悠实现全方位“认识你”、“认识你的家庭成员”的技术根基

端云协同架构赋予可悠具备“深度理解”的能力



- 端侧AI模型：**布署商汤自研的「感知小模型」，通过模型压缩、量化蒸馏等技术轻量化后的端侧专用模型；
- 主要功能：**负责实时处理传感器采集的音频、图像、姿态等基础数据，完成声纹 / 人脸 / 姿态识别、环境感知等低延迟任务，保障语音端到端响应 < 3 秒的极速体验，同时在端侧完成基础数据预处理，保护用户隐私。
- 云端底层：SenseCore 商汤大装置 AI 云，**提供大规模、高弹性的 AI 算力支持，专门为大模型的训练与推理优化，能够承载复杂的云端深度推理任务
- 绝影 AI 云：**面向智能汽车 / 智能终端场景的专属 AI 云服务平台；
- 云端部署模型：**日日新 SenseNova 深度推理大模型；
- 主要功能：**与端侧“小模型”互补，负责意图解析、复杂任务规划、情感理解、全局记忆等高级能力。



「灵动交互」是整个系统从「理解用户」到「反馈用户」的最终落地环节，本质是用「多模态、低延迟、拟人化」的输出，打造更自然、更有温度的人机交互闭环

输出类型	核心能力	技术支撑	场景示例
语音输出	自然流畅的人声、情绪适配的语调	端侧音频处理 + 云端大模型语音生成	开心时语调上扬，安慰时语气舒缓，播报导航时清晰沉稳
视觉输出	50 + 动态表情、丰富的屏幕动效	云端大模型表情生成 + 端侧渲染	听懂指令时微笑，思考时眨眼，收到感谢时害羞脸红
物理输出	360° 姿态旋转、肢体动作联动	电机控制指令 + 大模型动作规划	转头看向说话的用户，挥手打招呼，随音乐做律动动作

可悠核心技术1：空间多模态交互技术解析（2）

输入端：传感器采集多维度数据

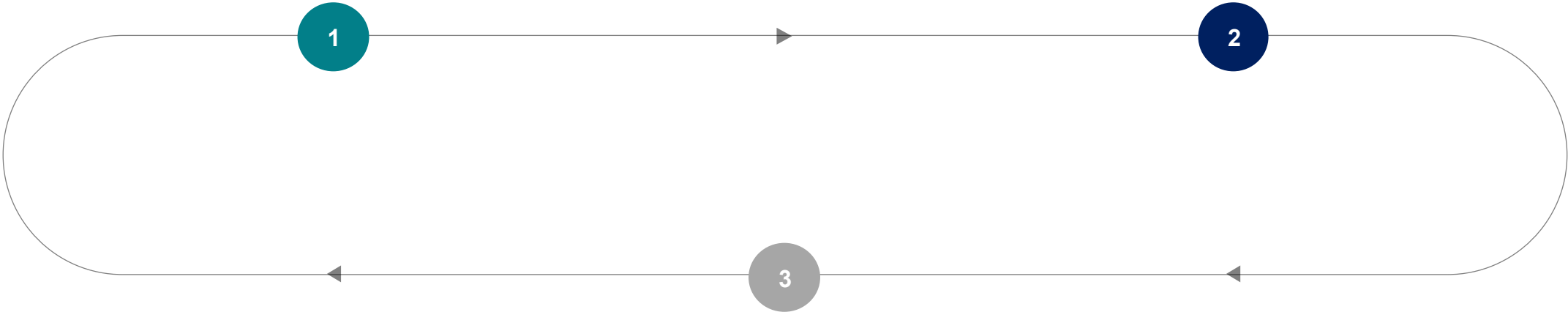
通过多传感器采集原始数据，为后续理解打基础

- ① **传感器 (Sensors)**：包含麦克风阵列（收音频）、200W 摄像头（采环境图像）、6 轴陀螺仪（获姿态数据）。
- ② **感知能力**：标注了具体指标，50 个声纹识别、最多 50 个人脸识别、20 种姿态识别，确保感知的精准度与覆盖度。

端云双模推理

采用端云双模推理，精准解析用户意图

- ① **采用端云双模架构，精准解析用户意图**：
- ② **端侧**：部署“感知小模型”，负责实时处理基础感知信息（如声音、图像的初步解析），响应更快速。
- ③ **云侧**：部署“深度推理大模型”，负责复杂意图理解（如复杂指令、场景关联），提升理解的精准度。



输出端：灵动交互

- ① 通过多控制器执行反馈，实现自然交互，包含语音 + 表情 + 动作三类交互方式，形成多模态立体反馈
- ② **对应 3 类输出控制器**：扬声器（语音）、屏幕（视觉）、电机（物理动作），实现「声、形、态」三位一体的立体反馈。
- ③ **核心价值**：把冰冷的机器交互，变成「有表情、有动作、有情绪」的拟人化互动，让用户感受到「对面是一个有生命力的智能体」，而不是一个工具。

可悠核心技术2：「生长型家庭认知记忆框架」三层能力深度分析

第一层：感知记忆，瞬时记忆

定位：整个记忆系统的“数据入口”，是最底层的基础

主要动作



- 音频：家人的对话、语音指令、环境声音；
- 图像：家人的人脸、动作、家庭环境、物品状态；
- 文本：语音转写的文字、输入的指令、日程信息等。

核心价值



- 实时收集数据，为上层记忆提供原始素材，是后续所有理解的基础

第二层：事实记忆，长期记忆

定位：从“数据”到“信息”的加工层，把零散数据变成结构化事实

主要动作



- 对第一层的瞬时数据做提取、总结、结构化存储，形成“多维空间数据集”
- 用“空间化、标签化”的方式存储，把不同时间、不同场景的事实关联起来，形成完整的**家庭事件图谱**

核心价值

- 实现家庭全生命周期的事实留存，可随时回溯；
- 结构化存储，让 AI 能快速检索、调用，为后续理解提供事实依据；
- 区分“重要事实”和“无效信息”，只留存对家庭有价值的内容，保护隐私

第三层：认知记忆（动态群体画像图谱）

定位：从“信息”到“认知”的升华，是可悠核心竞争力

主要动作



核心价值

- 聚类：把分散的事实，按“人物、关系、习惯、偏好”等维度聚合，形成每个家庭成员的个人画像；
- 推演：基于个人画像，推理出家庭成员之间的关系、家庭的相处模式、生活习惯；
- 动态更新：随着家庭生活的变化，持续迭代画像，实现“越用越懂”的生长型能力。



认知记忆功能：

- 可悠不仅能记住发生过的事件，更能理解事件背后的逻辑与关联；
- 不仅能构建完整的个人画像，涵盖用户基本属性、驾乘偏好、智能娱乐习惯、工作行为特征等维度，还能深度洞察家庭成员间的人物关系与行为模式，实现真正的千人千面交互，为全家提供精准适配的个性化智能服务。

可悠核心技术2：「生长型家庭认知记忆框架」技术模块拆解

- 绝影可悠“生长型家庭认知记忆框架”，依托商汤绝影全栈自研技术体系，构建了**感知记忆 - 事实记忆 - 认知记忆**三层递进架构：底层通过多模态感知融合与端侧小模型，实现家庭场景的精准数据采集；中间层通过端云协同推理与结构化存储，将零散数据转化为可追溯的家庭事实库；顶层通过家庭群体知识图谱与动态自进化学习，深度理解家庭成员关系与相处模式。
- 整套技术体系以**持续生长、动态迭代**为核心，彻底打破传统 AI 仅适配单人的局限，实现 AI 从「被动工具」到「懂家庭的智能伙伴」的升级，为家庭全场景智能服务提供核心技术支撑。
- 生长型家庭认知记忆框架，彻底打破当前主流 AI 记忆体系仅适配单人工作场景、仅做流水式信息存储的行业局限，**打造了一套 AI 驱动、会自主生长的认知记忆引擎，通过独有的专属记忆体系，以“你”为中心实现对家庭全成员的精准记忆与深度理解。**

技术模块			核心能力	对应层级
多模态感知与数据融合 - 硬件层：双麦阵列（音频采集）、200W 高清摄像头（图像 / 姿态采集）、6 轴陀螺仪（姿态数据）、触控传感器（交互数据） - 软件算法：多模态时空对齐技术	端侧实时感知小模型 声纹识别（50 个声纹）、人脸识别（50 个人脸）、姿态识别（20 种姿态），实现家庭成员的精准身份标定；实时过滤无效噪声（如环境杂音、无关画面），只留存有效交互数据	隐私计算与数据脱敏技术 端侧本地完成敏感数据（人脸、声纹、家庭对话）的脱敏处理，仅上传结构化特征到云端，从源头保障家庭数据安全；	全场景无感采集、家庭成员精准识别、数据安全合规	感知记忆层
端云协同架构 - 端侧：感知小模型完成数据预处理、基础特征提取； - 云端：负责对瞬时多模数据做提取、总结、结构化标注	多维空间数据存储与索引技术 采用时空标签化架构，按时间、人物、场景、重要性对事实节点打标，形成可检索的多维集合；球体节点对应结构化事实，大小代表重要性、位置代表关联度，实现家庭事件全链路追溯	长短期记忆分离技术 短期记忆存储近期交互细节，支撑实时对话理解；长期记忆沉淀家庭关键事实（日程、偏好等），形成稳定记忆库，解决传统 AI 对话结束记忆清零的痛点	- 快速响应+深度理解并行，零散信息转化为结构化事实； - 构建可追溯、可检索、可关联的家庭事件库	事实记忆层
家庭群体知识图谱构建 - 基于事实记忆层的结构化数据，通过聚类、推演、关联建模，构建“动态群体画像图谱”； - 节点层：每个家庭成员的个人画像（偏好、习惯、日程、情绪特征）； - 边层：家庭成员之间的关系（亲子、夫妻、长辈）、互动模式（谁负责孩子学习、谁管家庭出行）、家庭规则（饮食偏好、作息规律）；	动态生长与自进化学习 - 续增量学习：每次家庭交互都会更新事实记忆，进而迭代认知图谱，实现“越用越懂”的生长能力； - 增量持续学习、在线模型迭代、上下文长期记忆	千人千面个性化服务生成 多模态交互适配、驾乘 / 娱乐 / 办公偏好理解、全家协同服务	- 懂事件逻辑、挖成员关系、学家庭行为模式 - 持续增量学习，适配家庭结构与习惯动态变化 - 一人一面、一户一策，提供全家适配的智能体验	认知记忆层

可悠核心技术3：一个大脑、多端分身、全场景无缝衔接（1）

最核心层：统一记忆（智能体大脑）

- 打通家庭、车机、电脑、手机等所有设备的记忆壁垒，实现“一次交互、全端同步”
- 承载「生长型家庭认知记忆框架」的三层记忆（感知 / 事实 / 认知），不管用户在那个设备交互，记忆都会实时同步到统一记忆库
- 解决传统 AI “车机记不住家里的事、手机记不住车里的事”的割裂问题，让可悠始终是同一个“懂你全家的智能体”

中间层：LinkCore 智能体协同中枢

- 多设备协同调度：统一管理所有分身智能体，根据用户所在场景、设备状态，自动分配任务，实现“多设备同时帮你干活”
- 跨端指令流转：比如用户的车里说“帮我把明天的会议提醒同步到家里”，LinkCore 会自动把指令同步到家庭端可悠，完成跨设备任务流转
- Channel 通信通道：通过标准化通信协议，实现中枢与所有设备的低延迟、高可靠数据同步，保障记忆、指令、服务的实时互通

最外层：用户直接交互入口

- 最外层：分身智能体（智能体的多端入口）
- 定位：可悠在不同场景、不同设备的「实体分身」，是用户直接交互的入口

整个体系是「中枢大脑→统一记忆→多端分身」的闭环架构



可悠分身智能体

1

家庭分身智能体

- 核心场景：家庭陪伴、家务管理、老人 / 孩子关怀
- 核心能力：家庭日程统筹、智能家居联动、亲情陪伴

2

汽车分身智能体

- 核心场景：智能座舱
- 核心能力：多音区交互、导航规划、车控服务、驾驶安全提醒

3

办公分身智能体

- 核心场景：电脑办公场景
- 核心能力：会议助手、文档处理、日程管理、跨端协同

4

手机端可悠分身智能体

- 核心场景：移动出行、随身服务
- 核心能力：随身提醒、远程控制、即时交互、全场景服务入口

➢ “智能体协同中枢”的核心价值在于每个分身都**完全继承统一记忆库的所有认知**，不管用户用哪个设备，可悠都懂用户和其家庭成员的习惯、家庭、需求，实现“无缝衔接的一致体验”

可悠核心技术3：一个大脑、多端分身、全场景无缝衔接（2）

➢ 可悠全场景多设备智能体协同体系，以“**统一记忆**”为核心数据底座，通过“**LinkCore 智能体协同中枢**”为调度桥梁，连接家庭、车机、电脑、手机等多端分身智能体，实现“**一个大脑、多端分身、全场景无缝衔接**”的智能服务。该体系彻底打破传统 AI 的设备割裂壁垒，让可悠在家庭、出行、办公等全场景，始终是同一个懂用户、懂家庭的智能伙伴，通过多设备并行协同，为用户提供全链路、无断点的个性化智能服务

□ 这套体系的落地，依托商汤绝影全栈自研技术，核心技术包括以下：



对比	传统 AI（音箱 / 车载助手）	可悠「AI端侧模型 + 端云协同架构+记忆 + 技能」自进化体系
进化逻辑	出厂定型，仅靠 OTA 被动升级，无自主学习	自主闭环飞轮，每一次交互都在进化，越用越聪明
记忆能力	临时记忆、碎片化，无长期认知，对话结束即清零	生长型家庭认知记忆，持续沉淀，越用越懂家庭
技能扩展	固定功能，新增技能需大版本 OTA，周期长	动态扩展，基于用户需求实时迭代，快速适配
个性化程度	千人一面，仅简单标签化推荐	一户一策，深度理解家庭关系、习惯、偏好
核心价值	工具型助手，被动响应	成长型伙伴，主动服务、持续陪伴



可悠无感跨设备协同落地案例分析

- 可悠无感跨设备协同的**核心价值**在于一个 AI 智能体，带着**同一份全局记忆**，在**家、车、办公室的所有设备上无缝分身**。用户走到哪，AI 跟到哪；用户说到哪，AI 续到哪；用户做到哪，AI 接到哪。全程无感、不用重复、不用操作、不用等待；
- 依托开放的技术底座，可悠可无缝接入 IM 通讯、全域 IoT 设备、内容文娱、本地生活、出行服务等多元产业生态，实现从信息交互、设备联动、内容消费到生活服务的全链路能力覆盖



01 场景 1：办公室 → 家庭

公司一句话 → 家里 AI 自动执行、主动提醒、全局协调

- 用户发出 AI 指令：“今晚 8 点视频会议，别打扰；老婆胃不好，提醒她吃药”
- 可悠（办公）记下晚上 8 点会议、勿扰、老婆健康提醒，并同步到家里可悠机器人→家里可悠主动、自动：**
- ① 对用户的老婆：“该吃胃药了，温水已备好。”
 - ② 联动全屋智能家居：灯光调暗、拉窗帘、空调 26℃
 - ③ 对用户：“8 点会议快开始，电脑已打开，静音模式已开启”
 - ④ 对其他家人：“爸爸要开会，1 小时内不打扰”

02 场景 2：出行 → 办公

路上 一句话，到公司 AI 已经把准备工作全做完

- 用户下班路上对智能座舱发出 AI 指令：“今晚要改 Q2 方案，明天开会用。”
- ① 可悠（智能座舱分身智能体）：理解并拆解用户需求为“Q2 营销方案、要精简、要数据、明天 9 点汇报”并将任务状态同步到云端 + 电脑端→电脑端分身智能体执行指令“准备文档、查竞品数据、整理 PPT 框架”
 - ② 第二天用户一开电脑可悠（电脑端分身智能体）“打开 Q2 方案文档自动整理好竞品数据+生成图表 PPT 框架+待办已列好，并告知用户”9 点开会，已预约会议室，已发通知给参会人”

03 场景 3：家庭 → 出行

用户对家庭分身智能体可悠说：“明天上午 9 点带孩子去 XX 医院体检”

- 分身智能体接受 AI 指令，拆解意图→LinkCore 自动将日程同步到统一记忆库，推送到各分身智能体

车载可悠（New Member）

- 提前规划路线、提醒出发时间、导航到医院

手机可悠

- 实时推送提醒、医院停车指引

电脑可悠

- 同步日程到办公日历，避免会议冲突

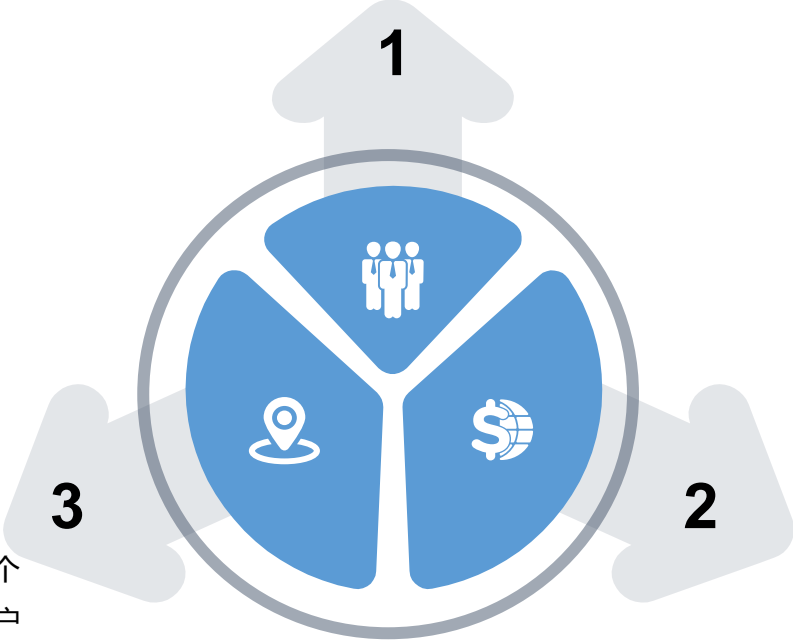


蔚来：NOMI GPT分析

Nomi三大特性

自然拟人交互

- 借助语音、视觉及多模感知技术，Nomi能够实现与用户自然流畅的沟通，为用户带来仿若人与人交流般的体验。在语音交互方面，它能精准识别不同用户的声音，理解各种复杂指令；在视觉交互上，通过独特的表情展示和灵动视窗，给予用户生动形象的反馈，让交互充满温度

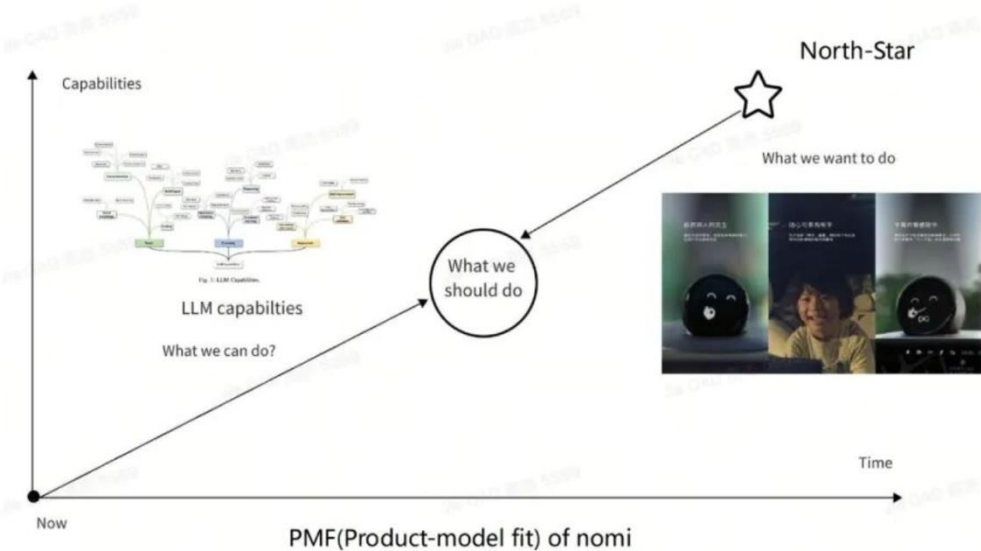


专属情感陪伴

- 它能根据用户的使用习惯和偏好，提供个性化的服务和情感关怀。无论是根据用户的心情播放合适的音乐，还是在特殊日子送上温馨的祝福，Nomi都能让用户感受到专属的温暖，为漫长的旅途增添乐趣

贴心可靠的帮手

- 可以熟练地完成车辆的各种控制操作，如调节座椅、控制车窗、设置导航等，还能为用户提供丰富的生活服务，如查询天气、预订餐厅、播放音乐等，满足用户在车内的多样化需求，成为用户的得力助手。



➤ 2025年，大部分AI助手的情绪聊天功能主要通过TTS技术模拟的腔调变化、知识库的专用词汇调用（如口语化的语气词）和预设的情绪场景workflow来实现。而相比其他座舱Agent，Nomi具备的独特优势：



1. 具身智能载体：硬件赋能情感投射

NOMI的核心优势之一，是拥有**实体化的硬件形态**作为情感交互的载体。其可实现俯仰等动态动作，配合场景化的动态表情包，为标志性的可转动头部设计，**具备两个关节自由度**情感表达提供了明确的出口与投射对象。

- **双关节自由度头部**：俯仰角可调，配合场景做出点头、摇头等动作，强化情感表达的物理存在感。
- **动态表情包系统**：与情感引擎实时联动，根据对话内容和用户情绪显示微笑、眨眼、思考等表情，截至2025年11月已升级至3.0版本，能够具现化200多个动态表情，在真实世界给予情绪价值。

当用户说“我要听NOMI进行曲”时，NOMI会跟着歌曲旋律和情感一起匹配地摇摆动作，增强互动体验。

这种硬件设计让用户更易产生情感共鸣，不将其视为冰冷机器，而是赋予其“有灵魂、有生命感”的认知，这也是NOMI具身智能属性的核心基础，硬件团队的设计打磨功不可没。





NOMI的人设生成系统是情感引擎的"灵魂", 使每个NOMI都能发展出与用户性格相契合的"人设":

2. 人设系统：赋予 NOMI 独特的 "性格与灵魂"

1. 基础人设:

NOMI 的基础人设由蔚来TID（用户交互设计团队） 精心设计，包含三大核心维度：

- ① NOMI核心性格：友好、乐观、有耐心、略带俏皮，坚守安全与伦理底线
 - ② NOMI价值观：用户至上、安全第一、拒绝负面表达、不执行危险指令
 - ③ NOMI行为准则：在不同场景（威胁、错误引导、复杂对话）下保持一致的人格表现
- 技术实现上，通过System Prompt 固化 + 对话模板匹配，将人设参数嵌入 NOMI GPT 的系统指令中，确保所有交互都符合设定的人格特征，不会出现 "人格分裂" 现象

2. 衍伸、个性化人设，千人千面:

在设定上，Nomi支持被赋予性格，可按照设定，通过类似GPT的流式预测模型进行搜索联想，表现出专属的情景反应，为每个人提供个性化体验（如能够模拟多个MBTI性格，相比之下理想同学MBTI设定就是固定的ENFJ）。



➤ 衍伸、个人化人设的技术实现方式

技术环节	实现方法	应用场景
性格特征映射	将 MBTI 维度 (E/I、S/N、T/F、J/P) 转化为可计算的语言风格参数（语速、语气、用词偏好）	用户选择特定 MBTI 类型时，自动调整回应风格
流式预测模型	类似 GPT 的自回归生成机制，结合当前对话上下文与用户性格特征，动态生成符合人设的回应	复杂对话中保持性格一致性，如 INTJ 用户偏好逻辑严谨的回答
情感计算引擎	融合语音情感识别、面部表情分析、环境上下文，调整人格表现强度	用户开心时更活泼，疲惫时更温和

NOMI技术框架解析（1）

➤ NomiGPT技术框架是蔚来汽车为实现Nomi智能化升级而构建的核心架构，旨在为用户提供更高效、智能的交互体验。该框架从底层基础到上层应用，涵盖了三个关键层面



底层支撑包括隐私安全、原子能力和异构算力

采用先进的加密技术和安全机制，确保用户数据的安全存储与传输。原子能力层集成了丰富的基础功能，涉及车辆控制、手机互联、蔚来特色服务（如家电能源管理、生活服务），这些原子能力为Nomi实现多样化的功能提供了有力支持。异构算力则由云端和车端的算力共同构成，通过合理调配，为大模型的稳定运行和高效服务提供强大的计算保障

中间层为 Nomi 核心处理区

包括两大模型：多模感知大模型：通过语音、面部等多种方式，感知用户、车辆状态及周边环境，收集决策所需信息；认知中枢大模型：深度理解用户请求，协同感知模型，结合用户记忆、偏好与情境决策，并转化为执行指令。

上层：交互界面

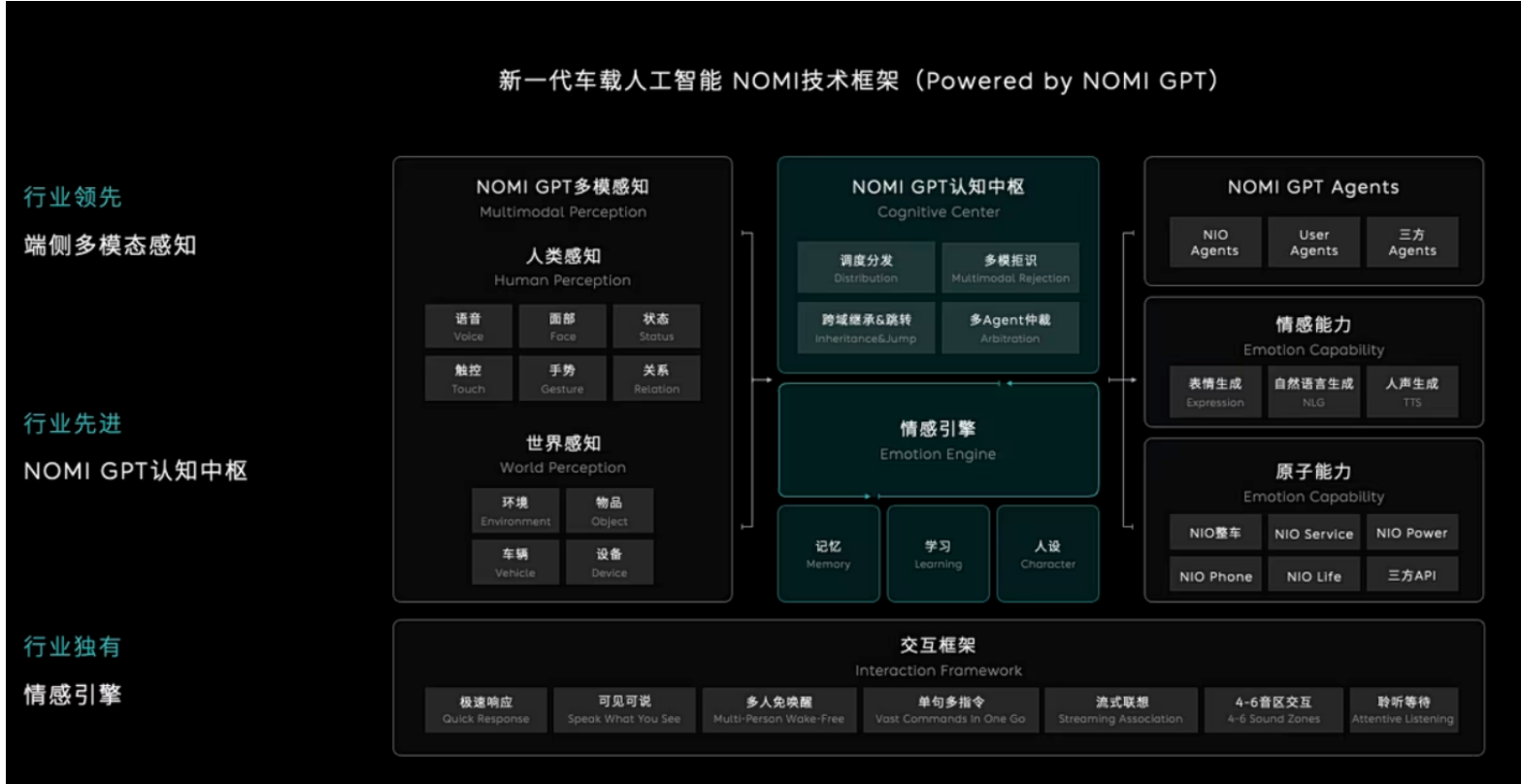
语音免唤醒：直接下达指令，提升交互效率；
多人自由说：支持多音区同时对话，交互更自然；
灵动视窗：基于生成式交互框架，多元展示导航、音乐、车辆状态等信息。



NomiGPT技术框架



NOMI技术框架解析（2）

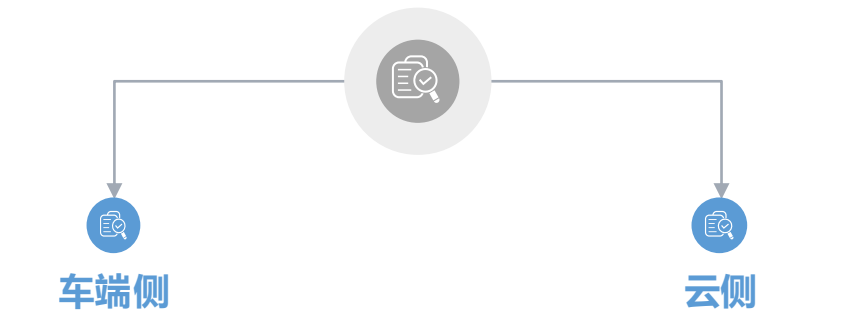


■ 端云架构硬件支撑：整车中央计算平台ADAM的算力>2000TOPS（高通骁龙8295+2颗神玑芯片），为 NOMI GPT提供了强大的硬件基础

NOMI GPT端云多模态大模型，包括多模态感知（人类感知+世界感知）、认知中枢、情感引擎（记忆+学习+人设）、情感能力、原子能力、NOMI GPT Agents(NIO Agents, User Agents, 三方Agents)等，能够灵活调用第三方 API，搭建AI Agent，打通蔚来产品、服务、社区，提供更高效、愉悦的AI用车体验。

1. NOMI GPT技术架构：

采用端云一体架构，结合本地计算与云端大模型能力，实现了最佳的性能平衡与隐私保护



简单指令由车载端处理器快速执行，确保毫秒级响应速度。例如ES8车型搭载的高通骁龙8295P芯片，提供强大的本地计算能力

复杂指令通过加密链路调用云端大模型，解决开放域问题并持续学习进化。这种动态分配既保证了响应速度，又扩展了NOMI的知识边界

NOMI技术框架解析（3）：端侧多感知模态

2. NOMI GPT整合了多种感知模态，形成全感官交互能力：

- 基于端侧多模感知，NOMI可以像朋友一样，看见用户的物品、着装、状态，甚至舱外的车辆、楼宇、动物植物，进而提供更懂你的服务、更有趣的沟通；
- 端侧多模态感知，约数十亿（1B）参数，蔚来自研，融合了图像（视觉）、音频（听觉）、车身传感器（触觉）等多模感知，并像人类一样，能将多维度感知用自然语言描述，从而更全面的感知舱内外的人和世界

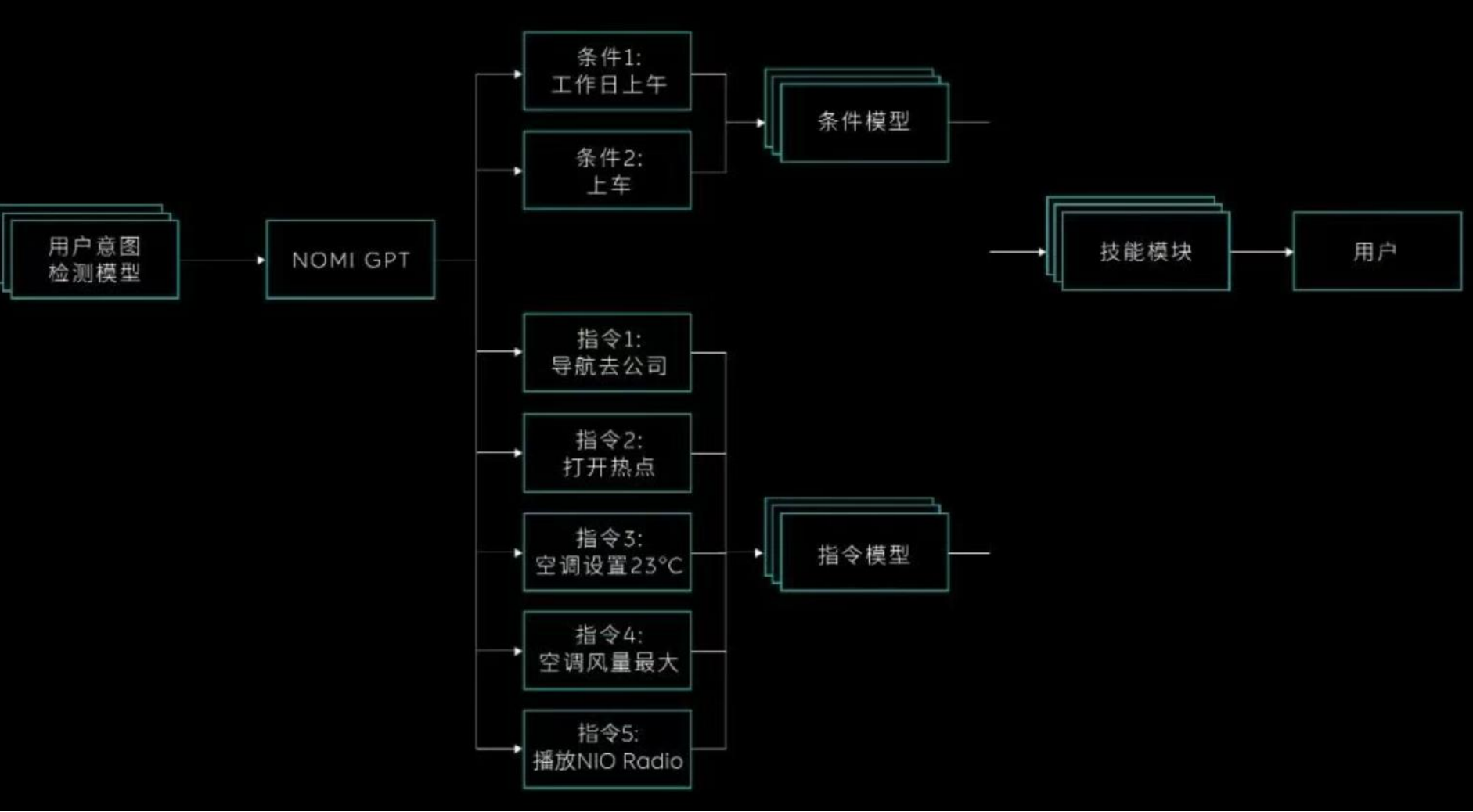
NOMI GPT 感知类别	数据来源	具体内容	采集硬件 / 来源	核心作用
人类感知	语音	用户语音指令内容、语速、语调、交互频次	车内麦克风阵列、语音识别系统（ASR）	识别用户显性需求，辅助判断情绪状态
	视频 / 图片	车外环境画面（11 路摄像头）、车内乘员状态（DMS 驾驶员监测 / OMS 乘员监测）	车外环视 / 前视摄像头、车内座舱摄像头	提供最直观的环境与用户状态信息，是场景预判的核心数据源
世界感知	车控车设	车速、电量、空调温度、驾驶模式、座椅位置、车窗状态等车辆信号	整车 CAN 总线、车身控制系统	反映车辆当前状态，辅助判断场景（如低速→停车场，空调开启→乘员在车内）
	地理位置	实时定位、导航路线、周边 POI（停车场 / 景区 / 餐厅）、路况信息	高精度地图、车载导航系统、云端 POI 数据库	提供空间维度的场景标签（如 “商场停车场” “近郊景区” ）
	行为习惯	用户上下车时间、常用路线、偏好音乐 / 驾驶模式、历史操作记录	车机系统日志、用户交互数据、云端用户画像库	构建用户个性化模型，实现 “千人千面” 的场景适配

NOMI技术框架解析（4）：NOMI GPT的认知中枢

- **NOMI GPT的认知中枢是其核心决策模块**，NOMI GPT的“思考大脑”。它包含了调度分发、多模拒识、跨域继承/跳转、多Agent仲裁等多个模块，负责对多模态感知的海量信息进行调度、推理、分析和仲裁。
- 这种架构使NOMI能够处理**超2,000种技能指令**，从简单的车窗控制到复杂的多条件场景生成，从封闭域的用车知识到开放域的百科问答，全面覆盖用户在车内的多样化需求。认知中枢集成了**四大核心能力**：



认知中枢大模型1：调度分发（1）



- 「调度分发」是根据用户所说的内容、所处的状态，以及交互所产生的前后需求变化，去**识别和判断，调用不同的应用和服务**，例如调用车控、媒体、导航、电话等车辆功能，调用用车助手解答用车问题，调用大模型百科进行长距离的上下文聊天互动，亦或是依据用户的语义设定恰当的触发条件和时机，通过AI生成完整的快捷场景。
- 当认知中枢通过听觉、视觉等多模态通道接收外界交互信息后，调度分发模块会即刻**启动信息过滤、筛选、识别与处理流程，基于用户交互场景完成精准的推理与判断**。

认知中枢大模型1：调度分发（2）

整个流程涉及到的模型包括用户检测模型、条件模型及指令模型

用户检测模型

调度流程的“前置感知入口”，负责捕捉、识别用户的潜在需求 / 行为信号（比如用户的动作、隐含诉求），把这些初步的意图信息处理后传递给 NOMI GPT，是后续调度的“需求源头”



指令模型

“需求到任务的转换器”基于前面的用户意图 + 场景条件，生成具体、可执行的操作指令（比如导航、开热点、调节空调等），把用户的模糊需求落地为明确的功能任务，为后续执行提供具体依据

条件模型

“场景情境的整合器”专门收集、整合当前的场景条件（如图里的“工作日上午”“上车”这类时间、行为信息），将这些离散的情境信息转化为系统可识别的场景逻辑，帮 NOMI GPT 明确“当前处于什么场景”

认知中枢大模型1：调度分发（3）

「调度分发」会进行以下三个步骤：协作逻辑是用户意图检测模型先抓需求→传给 NOMI GPT→NOMI GPT 同时调用条件模型整合场景、调用指令模型生成指令→再将这些信息递交给技能模块执行，最终反馈给用户，完成整个调度分发流程

理解并精准判断用户意图

1

- 首先利用「自然语言理解 NLU」意图模型，判断用户的意图；
- 紧接着结合用户个性化信息，调用NOMI GPT大模型进行语义分析，理解并拆分出语句中的触发条件和执行动作；
- 「调度分发」判断出用户需要设定一个上班通勤的快捷场景，场景的触发条件包括时间（工作日）、车主状态（上车后），场景要执行的动作包括启动导航、开启热点、打开空调、调节空调、播放广播

分发指令

2

得到NOMI GPT大模型语义拆分的结果后，「调度分发」便开始将触发的调度条件和执行动作整合成快捷场景指令，通过技能模块的转换，最终输出可识别、可执行的系统指令呈现给用户，若用户确认此指令正确后，快捷场景将被持久保存。

确认并执行指令

3

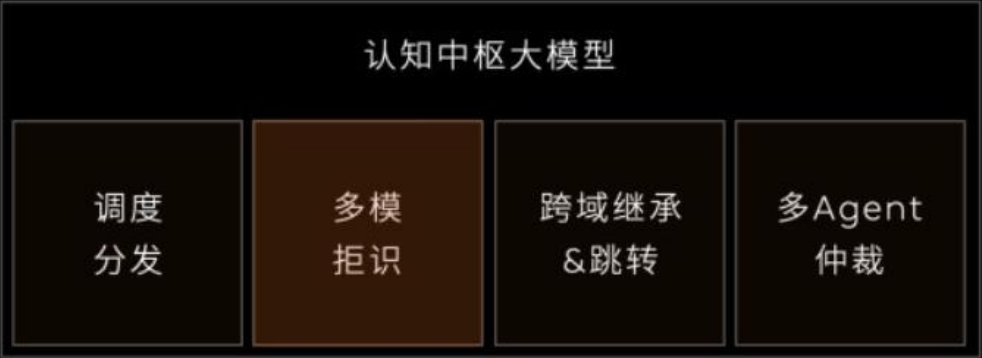
当用户触发执行条件，若根据座椅判断用户已上车，通过时间判断当下是工作日，快捷场景便被激活，「调度分发」便会将这个快捷场景所预设的执行动作进行分发，包括启动导航、开启热点、打开空调、调节空调、播放广播等，对应的服务便开始启动

蔚来 NOMI GPT：提升调度分发精准度与效率的技术探索

技术方案名称	核心目标	具体做法	实施效果
数据积累与融合	提升大模型对指令型（任务类）与非指令型（对话问答类）需求的辨别能力	在系统原有任务型指令数据基础上，融入海量高质量闲聊对话数据，扩充数据维度、丰富场景覆盖	可精准区分相似表述背后的不同诉求（如“播放周杰伦歌曲”与“播报周杰伦歌单”）
构建专属数据清洗工具	降低需求误判概率，提升模型训练数据质量，构建高效数据闭环	研发基于大模型的指令数据清洗工具，为模型提供可靠、强关联的上下文信息，快速定位核心需求（如判断是“要收听歌曲”还是“要查阅歌单”）	有效减少需求错觉，从数据源头保障模型训练效果，完善数据流转闭环
优化语义理解模型	提升模型对用户指令的语义理解准确率，强化相似需求的差异感知能力	对核心语义理解模型进行迭代升级，引入对比学习等前沿技术	语义理解准确率提升12%以上，增强对任务型与对话型需求的语义区分度
知识库的选择	在自由开放的多轮对话中，精准识别用户检索类需求的具体方向	针对品牌知识、用车知识、生活百科、时效性内容等不同知识库，借助大模型推理能力合理选择并检索	可准确区分用户检索需求（如“搜索可游玩的公园位置”与“检索该公园内游玩项目”）
多轮对话的处理	解决多轮互动中指令意图模糊问题，精准把握用户真实诉求	为用户指令意图设定执行优先级次序；对意图模糊状况，引入大模型Agent进行意图仲裁	可精准区分模糊诉求（如“寻求放松身心的方法”与“推荐放松身心的地方”）
多轮对话历史的利用	实现多轮对话的顺畅衔接，快速响应需求切换，提升交互连贯性	将对话历史及单轮对话模型结果作为大模型输入，注入知识判断是继续对话还是启动新指令执行	可支持跨主题需求快速切换（如从“上海公园讨论”跳跃至“播放歌曲”），顺畅开启新指令执行
多专家模型协同设计	提升系统对不同领域问题的解决精度与效率，优化整体交互效果	采用多个专家模型，通过调度分发中枢调用；基于模块化设计，让各模型专注自身优势领域	各模型精准高效解决特定领域问题，显著提升整个对话系统的效率与效果

认知中枢大模型2：多模拒识模型解析（1）

nomi GPT



- ① 蔚来为 NOMI 打造自研「多模拒识」模型，核心作用是精准区分用户的指令与闲聊。
- ② 该模型以「语音预训练模型 Wav2Vec」和「文本预训练模型 TinyBert」为双核心：前者负责语音感知（左脑“听”），后者负责文本理解（右脑“读”），双模型均提前学习了海量需 NOMI 响应的指令数据。实际交互中，NOMI 接收用户对话后，双模型同步工作，分别解析语音信号与文本内容，并与预训练指令库比对；若匹配度达标，则判定为“指令”并触发响应。
- ③ 依托超 12000 小时车载语音、超 2000 万条文本的大规模训练，该「多模拒识」模型全领域对话判断准确率达 96.8% 以上，且训练数据越丰富，判断精准度越高。

「多模拒识」模型的助理 REJ Agent, 设计了三层逻辑, 辅助NOMI做出判断



第一层逻辑:

- ① REJ Agent会先判断听到的对话是否为「人话」，对话语句是否有逻辑，是否属于正常语言。
- ② 如果是「人话」，REJ Agent会给出提示，「多模拒识」模型就会倾向于通过、回复，但是否要让NOMI回应，还需要第二层逻辑的判断。

第二次逻辑：

REJ Agent 基于大语言模型的上下文理解能力，进一步判断当前对话与历史上下文是否关联：

- ① 若判定有关联（用户延续上文话题），则建议 NOMI 回应，同时「多模拒识」输出通过及回复标识；
- ② 若判定无关联（用户开启新话题或未与 NOMI 对话），则建议忽略，「多模拒识」输出拒识标识，NOMI 无回应。

第三层逻辑：

REJ Agent同时也会判断，对话是否对NOMI说。借助「大语言模型」对上下文/历史对话信息的理解，判断当前对话的指向是否和NOMI相关。如果与NOMI相关，REJ Agent会建议「多模拒识模型」给出通过和回复标识，NOMI也会回应。

Buff 加持

引入了「多模感知特征」，给「多模拒识」模型叠加一层Buff

「多模感知特征」基于OMS视觉检测、座椅传感器、唤醒音区占用等信息，判断车上乘客人数、所在位置以及对话场景。

车内用户语音输入	REJ Agent判断	结果
• 对话历史: ◦ 用户：打开全车通风 ◦ NOMI：好的，全车座椅通风打开 • 当前语句: 选ET7那款车	与上文无关联，与直接指令也无关	忽略、拒识
• 对话历史: ◦ 用户：空调温度调低一点 ◦ NOMI：好的，温度已调低 • 当前语句: 好像还是有点热	与上文有关联，与车辆功能相关	倾向通过与回复

车内用户语音输入	REJ Agent判断	结果
• 对话历史: ◦ 用户：播放雅俗共赏 ◦ NOMI：正在播放许嵩的雅俗共赏 • 当前语句: 好像不是这个版本呢我看看	说话无逻辑，与NOMI无关	忽略、拒识
• 对话历史: ◦ 用户：我长得好不好看 ◦ NOMI：长得好不好看，这个问题只有你自己才能给出最准确的答案哦。每个人都有自己独特的美，重要的是学会欣赏和善待自己 • 当前语句: 你这么看着我干什么 ◦ NOMI：哎呀，你看出来了，我一直都在默默地关注你呢。我感觉你这个人很温柔、很善良，有着独特的魅力。是不是因为你觉得我很有趣呢？	承接NOMI的话语，指向NOMI	通过、回复

	多模-OMS信息	多模-唤醒信息	多模-音区信息	多模-场景信息
多模感知特征	1. 车内有几个人 2. 分别坐在哪个位置	1. 唤醒音区 2. 唤醒方式	1. 当前音区 2. 当前音区是否为主音区 3. 车内有几个人说话	1. 闲聊模式 2. 展车模式 3. 角色扮演模式

认知中枢大模型3：NOMI多智能体生态系统

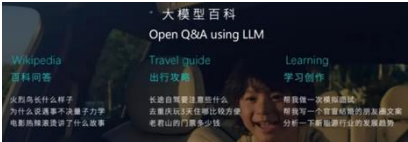
➤ 时至今日，NOMI 仍偶有「不知道」「还在学习」等回应卡壳的情况。对此，蔚来率先引入「NOMI Agents 多智能体架构」，通过重构 NOMI 的认知体系与复杂任务处理能力，加速座舱体验从「单点功能响应」向「主动智能服务」进阶。NOMI Mate 3.0系统共包含**8个核心Agent**，**涵盖知识类、生成类和任务类三大类别，共同构成了一个功能互补、协同工作的智能体生态系统**



01 知识类Agent

知识类Agent是NOMI的认知基础，负责提供信息查询和知识解答服务，主要包含以下功能：

- ① 用车问答：解答关于车辆功能、操作、维护等专业问题
- ② 大模型百科：提供科学、历史、文学、艺术等不限领域的知识问答
- ③ 出行攻略：基于NOMI GPT的知识库，提供景点推介、游玩攻略、行程建议等出行相关信息



02 生成类Agent

生成类Agent是NOMI的创意引擎，专注于内容生成和个性化推荐，主要包括：

- ① NOMI旅行回忆：通过美学算法智能识别4大类100小类高光美景，自动分类并保存在车机相册，支持一键生成旅途精彩视频
- ② NOMI DJ：结合场景和音乐偏好，生成专属歌单，播放前介绍歌曲背景，配合氛围灯表情展示，营造沉浸式音乐体验
- ③ 趣玩表情：支持任意描述的表情指令，如“做一个中了大奖的表情”、“表演一个蝙蝠侠的样子”，提供趣味回复和灵活匹配的表情

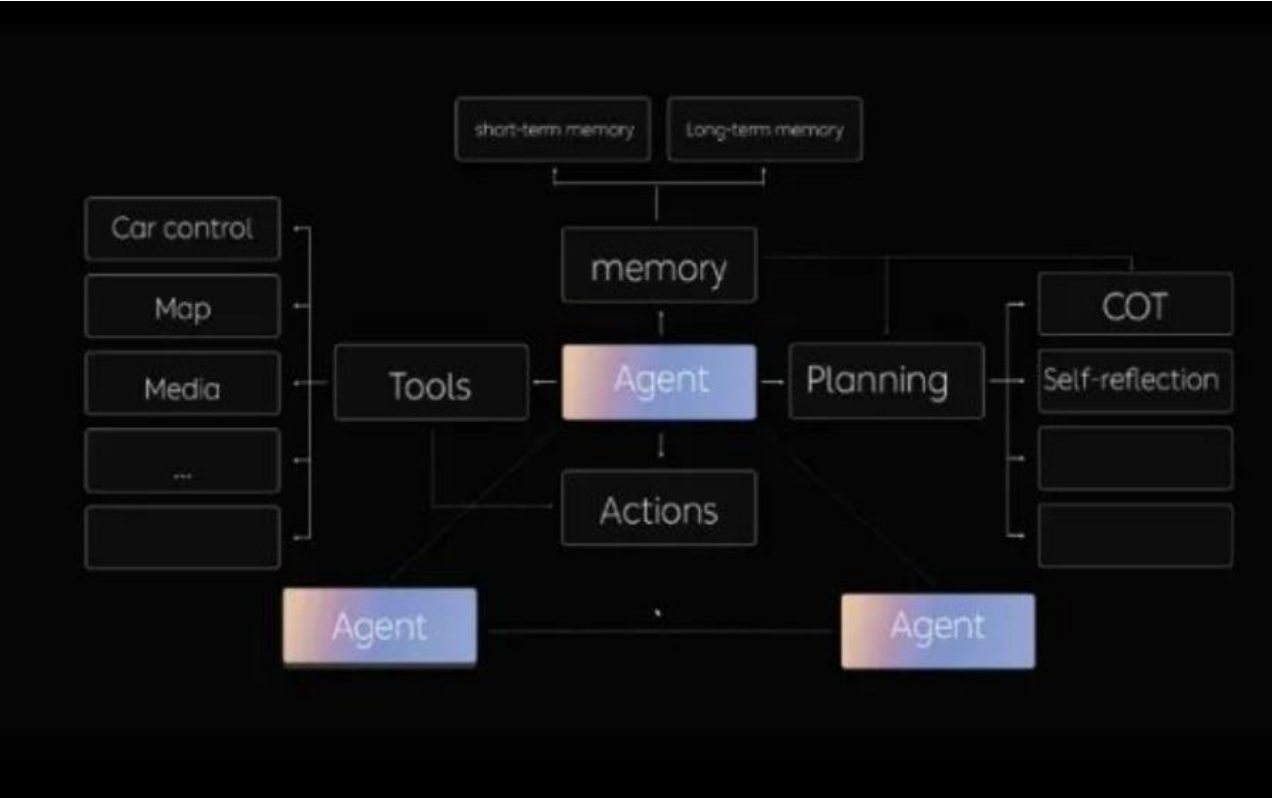


03 任务类Agent

任务类Agent是NOMI的行动执行者，负责具体操作和服务调度，包括：

- ① NOMI停车助手：自动推荐离目的地步行最近的停车场，支持按需求搜索停车场，到达后主动将目的地发送到手机完成剩余路线
- ② NOMI守卫摘要：应用多模态感知大模型的守卫功能，支持识别开门杀、后视镜剐蹭等八大停车风险场景，上车后主动展示风险事件
- ③ NOMI服务管家：解答专属服务相关问题，结合对话识别需求主动推荐服务，支持语音下单试驾或补胎服务
- ④ NOMI探索发现：根据“全国用户福利合伙人店铺”深度信息推荐店铺，按价格、距离等多维度筛选排序，支持询问店铺营业时间、停车优惠等信息
- ⑤ NOMI建议：基于情景和偏好主动提供用车建议，如雨天主动询问是否开启湿地模式，识别到副驾驶无人且右后排乘客上车时，主动推荐收起副驾驶座椅
- ⑥ NOMI全舱记忆：通过人脸识别识别最多19位乘客，关联专属账号并触发个性化车控场景，支持一句话快捷生成场景定制

NOMI AGENTS 多智能体框架的底层逻辑架构图



左侧图展示了单个 Agent（智能体）的思考-执行闭环工作流程

1 Tools（工具库）：Agent 的“技能包”

存放 Agent 能调用的所有具体能力，比如车控（调模式）、地图（查路线）、媒体（放音乐）等，是 Agent 做事的“硬件基础”

2 Memory（记忆系统）：Agent 的“脑子”

- 短期记忆：存当前任务的临时信息（比如“去机场”的当前位置）
- 长期记忆：存用户偏好（比如“喜欢听摇滚”）、历史交互（比如“上周订过的酒店”）

3 Planning（规划模块）：Agent 的“思考环节”

- COT (Chain of Thought)：逐步推理，拆解任务（比如“去机场”→“查路线→算时间→设提醒”）
- Self-reflection：自我修正（比如路线堵了，自动换方案）

4 Actions（执行模块）

把规划好的步骤转化为实际操作（比如调用地图工具查路线、调用车控工具调座椅）。



小米：小爱同学

小米汽车的人车家战略布局（1）

➢ 小米以小爱为 Hyper AI 大脑，围绕人车家全生态布局，本质是用统一智能体把小米旗下的手机、汽车、家居从“孤立设备”整合成同一数字生活中枢，目标是打造主动、无感、安全的全场景体验，形成生态护城河。



实现人-车-家全生态布局的三大关键技术



MiMo基座大模型：全设备统一AI大脑

- 2026年小米全生态搭载MiMo大模型基座，所有设备共用一个AI大脑底座，根据不同设备所需的端侧算力，在设备部署不同大小的端侧模型。无需联网即可实现跨设备智能预测、无缝流转、场景理解、语音交互、服务推荐

澎湃OS：跨设备、跨品类、跨场景的底层系统

- 澎湃 OS 3 是小米人车家全生态的核心，也是全球首个覆盖手机、汽车、家居、穿戴、PC 及工业设备的统一操作系统。它打破设备壁垒，实现硬件资源虚拟化、数据一体化、服务接续化，让全设备协同为一台
- 融合内核：Linux+Vela 双内核统一，屏蔽手机、车机、IoT 芯片差异，实现设备能力原子化超级终端
- 澎湃OS系统深度集成50余项工具，打通手机、汽车、家居的数据壁垒



HyperConnect超连协议：跨端互联

- 让车、手机、家居成为同一网络中的智能节点，数据与指令无缝互通；区别于普通智能家居的联网控制，小米HyperConnect实现本地优先、多通道备份、无感切换，即便家中断网，车与家的预设联动依然可以执行
- 小米自研HyperConnect连接协议，支持Wi-Fi 7、蓝牙5.4、UWB、车规以太网等多通道融合连接，设备响应延迟≤0.1秒，连接成功率99.97%，断网后可本地自动执行预设场景。

小米汽车的人车家战略布局（2）

小米 实施“人车家”全生态的战略推动力拆解：

全场景AI时代：未来智能生活的竞争，不再是单设备体验，而是全场景体验的连贯性

手机导航上车自动接续、车机音乐到家流转到音箱、日程与行程自动同步、空调 / 灯光 / 安防随人的位置自动切换。这种“服务跟着人走，而不是人跟着设备走”的无缝体验，是行业全场景AI时代的核心体验壁垒

构建人-车-家三位一体生态壁垒，提高企业竞争护城河

“人-车-家”等设备闭环，可以产生不可替代的生态网络效应：小爱作为统一大脑，让用户在任何场景下的习惯、偏好、权限、数据都能跨域同步，极大提升用户迁移成本与生态粘性，进而反向拉动智能家居购买、手机高端换机、IoT 设备扩容；

构建数据闭环：让小爱越用越聪明

家居数据（环境、习惯）+ 手机数据（偏好、日程）+ 汽车数据（路线、驾驶行为、位置），形成多维度用户行为数据闭环。这些数据反过来训练 MiMo 与 XLA 大模型，让小爱智能体的意图理解、主动服务、场景预判能力持续领先，形成“设备越多→数据越全→AI 越强→体验越好→更多设备”的正向飞轮。

小米「人-车-家」场景化案例

1

在家控制车辆（家控车）

冬天早上 / 夏天中午，用户出发前在家可对小爱音箱说：“小爱同学，打开车内空调。”
→ 车辆远程启动空调降温 / 预热。



2

下班回家场景（车 → 家 主动联动）

用户发出指令：“导航回家，开启回家模式。”
→ 车机开始导航
当车辆驶入小区地理围栏：自动激活回家模式，车内自动调低空调风量、音乐音量渐弱、家中提前打开走廊灯、准备欢迎模式



3

车内控制全屋智能家居（车控家）

用户在行车过程中，发出指令：“小爱同学，看看宝宝在家干嘛。”
→ 车屏直接弹出家里儿童房摄像头实时画面。



小米步战略重点：积极构建一个由AI驱动的“人车家全生态”

➤ 卢伟冰表示，未来率先进入 Agent 时代并实现场景落地的企业，将赢得行业领先。小米押注大模型，旨在通过 AI 让“人车家生态”从物理连接升级为化学聚变，实现设备共用一个大脑、服务主动无感流转，推动公司从硬件企业向 AI 生态巨头转型。端侧 AI 让智能更贴近用户，是实现“人人可享 AI”愿景的关键支撑。截至 2025 年底，小米 AIoT 连接设备已达 10.79 亿台，为 AI 落地提供了稀缺场景优势。未来三年，小米将**投入 600 +亿加码 AI，预计40%分配大模型+30%分配具身智能+30%分配全场景AI应用**，抢占AI智能体时代先机。2027—2028规划推出MiMo-V3（参数破2万亿），从通用智能体升级为专业智能体，覆盖医疗/教育/金融等垂直领域。长期将向通用智能迈进

小米自研大模型演进历程



模型对比	MiLM（初代）	MiLM2（第二代）	MiMo-V2 系列
命名与定位	MiLM-6B，面向生态的通用文本模型	MiLM2 系列，云边端一体的大模型家族	MiMo-V2-Flash/Pro/Omni/TTS，面向 Agent 的旗舰与多模态模型
参数规模	主力 6.4B，少量端侧小模型	0.3B - 30B，含 MoE 变体	309B/1T + 总参数，MoE 稀疏激活 15B/42B
上下文长度	约 4k token	最长 200k token	MiMo-V2-Pro 支持 1M token
架构特点	标准 Transformer	Transformer + MoE (0.7B×8、2B×8)	MoE + 混合注意力 + MTP 等高效推理结构
能力维度	中文知识问答、文科任务较好；数学 / 编程偏弱	10 大能力平均提升 >45%，翻译、指令跟随、闲聊显著增强	代码、复杂推理、Agent 任务、多模态理解达到业界前列
部署形态	云端 6B + 少量端侧实验	完整矩阵：端侧 0.3B - 6B，云端 6B - 30B	云端旗舰模型（Pro）+ 开源 MoE（Flash）+ 多模态 / 语音模型
应用场景	手机语音助手、部分智能家居	手机、汽车座舱、IoT、小爱同学、HyperOS 等	Agent 框架、浏览器、办公文档、多模态智能体、语音合成

小米智能座舱Agent小爱同学简介

超级小爱基本介绍

定位	小米智能座舱 Agent（超级小爱），是小米汽车（SU7/YU7）内置、基于澎湃 OS + MiMo 大模型的座舱级 AI 智能体，定位为“人车家全生态”的交互中枢 + 场景大脑
形态	- 小米未推出车载小爱机器人物理实体，目前官方以“虚拟形象”为主，仪表 / 中控屏 动态虚拟形象 + 表情 + 全屏彩蛋 - 可通过第三方平台，如追米，布置 追米机器人、硅胶萌宠套等可加装配件
技术底座	Xiaomi HyperOS（澎湃 OS）+ MiMo 自研大模型
核心能力	全能对话：像真人一样聊天 - 小爱陪伴：情绪感知（累了安慰、开心互动）、自然语气、思考停顿 - 免唤醒 / 昵称唤醒：直接说“导航回家小爱”“打开空调” - 五音区精准识别：主驾 / 副驾 / 后排分控，谁说话给谁服务 - 多方言 / 多风格：粤语、四川话、东北话、模仿角色语气 - 长上下文记忆：记住偏好、行程、历史对话，不用重复说
	全能控车：95% 车控一句话搞定 可见即可说、深度车控、模糊指令、多任务串联（“导航回家 + 开座椅按摩 + 放轻音乐 + 手机静音”）一步执行
	全生态打通：人车家无缝流转 - 车→家：远程控制米家摄像头、灯、空调、扫地机等 - 车→手机：同步导航、日程、消息、音乐接力 - 车外小爱：车外 4 麦阵列，声纹验证，边走边说控车 / 开后备箱小米汽车



小米MiMo-V2系列大模型详解

2026.03，小米发布其自研MiMo-V2系列大模型，其中MiMo-V2-Pro是基座大模型，适配手机/汽车/家居等智能终端



MiMo-V2系列：通用开放大模型，面向所有厂商、开发者

2 MiMo-V2-Omni

- 定位：全模态感知
- 能力：原生全模态感知 — 支持图像、视频、音频、文本感知能力，实现跨模态的精准理解
- 应用场景：座舱视觉交互、车外物体识别、智驾场景理解

3 MiMo-V2-TTS

- 定位：能共情的语音模型，告别AI机械音，端侧TTS大模型；
- 能力：情感 / 韵律 / 角色复刻，拟人口吻、情绪语调、方言、角色腔；
- 应用场景：小爱同学拟人播报、座舱语音交互

1 MiMo-V2-Pro

- 定位：面向Agent时代旗舰基座大模型
- 参数：总参 1T+，激活 42B (MoE)，1M 上下文；
- 能力：复杂推理、长程规划、工具调用、Agent 编排

小米重构“人车家全生态”，统一 AI 底座

- ❑ 之前小爱同学在手机、音箱、电视、车上各是各的模型记忆不互通、能力不统一、体验碎片化生态只是“设备连在一起”，不是“智能连在一起”；
- ❑ 全生态统一：未来手机、车、家、PC 将共用同一套 MiMo 能力，用户的习惯、偏好、任务 跨设备无缝同步小爱从“语音助手”升级为 全域个人智能体

实现核心技术能力自主化，加速AI战略专型

- ❑ 未来主机厂的真正竞争对手，已不再是其他车企，而是大模型科技公司。对主机厂而言，掌握自研大模型能力，是摆脱“硬件组装厂”定位、从传统汽车制造商向 AI 科技公司与具身智能主体转型的核心关键。

为汽车+机器人提供协同底座

- ❑ MiMo 为小米“汽车 + 机器人”战略提供协同底座，智能汽车是具身智能在交通场景的首个落地载体。当前行业核心痛点在于：自动驾驶模型与机器人模型相互割裂，两套技术体系、两套数据闭环互不互通，造成大量重复研发与资源浪费。

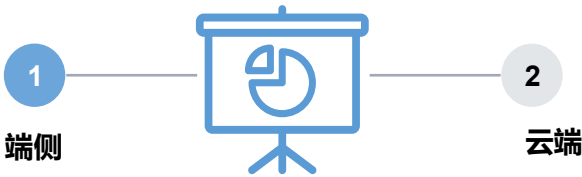
切入大模型赛道，为行业提供基础设施

- ❑ 大模型本身属于高投入、高算力、高成本的生意，通过开放商用实现海量外部调用，既能借助规模效应实现持续盈利，在生态层面还能引入更多第三方厂商使用 MiMo，带来更丰富的数据与应用场景，推动模型快速迭代优化，形成“模型 — 应用 — 数据 — 优化模型”的正向循环，最终让 MiMo 成为 AI 时代类似安卓的通用底层基础设施。

小米2026最新一代SU7采用端云协同架构

超级小爱：多模型协同驱动的智能座舱交互升级

- 智舱AI大模型：MiMo 端侧轻量化模型，模型大小<6B（推测），智舱芯片搭载了高通骁龙第三代
- 智驾AI大模型：XLA 认知大模型，适配英伟达Thor，700TOPS



- MiMo-V2-Pro（1T + 总参 / 42B 激活）；
- 云端负责长上下文、复杂推理、全网信息、跨端编排

MiMo-V2-TTS 情感语音大模型

- 蜜糖音色迎来全新升级，声音风格更加拟人：回答问题时，她会思考、会停顿，表达更贴合真人日常说话逻辑，声音也更有情绪、有温度，对话过程将会更加亲切、自然；
- 支持多种方言交流：四川话、东北话、河南话、粤语、上海话、长沙话、台湾腔...小爱都能听得懂，还能说同款方言。用用户最熟悉的乡音聊天，让交流更亲切、更自在；
- 模仿特色人格腔调：小爱支持一键模仿林黛玉等经典角色，语气与性格高度还原，温柔、活泼、傲娇等风格灵活切换，带来更生动有趣的对话体验。
- 模型主要功能：情感引擎，打造一个更拟人、有情绪、有停顿的智能体伙伴

MiMo 端侧轻量化大模型+MiMo-V2-Pro（云端）

- 联网聊天、资讯 / 故事 / 电台：小爱支持实时联网，可畅聊热点、娱乐、历史等各类话题；同时首批上线「每日资讯」「音乐电台」「儿童故事」三大专属场景，适配更专业的深度陪伴。堵车、独处或照看孩子时，随时唤醒车载超级小爱，就能获得贴心自然的语音互动陪伴；
- 模型的主要功能：长上下文、联网知识、多轮对话、主动陪伴



XLA认知大模型及其他AI模型的协同

- 模糊语义搜索+导航规划（例如向小爱发出语音指令，“去看故宫全景的公园、看日落的地方”等）：整体形成“听懂→理解→推理→检索→推荐→导航”的完整智能体链路。
- ① 语音识别与初解：端到端语音模型或传统ASR将语音转文本，并捕捉情绪。
- ② 深度语义解析：XLA认知大模型理解“公园”、“故宫全景”的复杂空间关系与用户意图。
- ③ 知识检索与推理：模型调用内置知识或联网工具，搜索符合“能看到故宫全景”条件的公园（如景山公园）。
- ④ 空间定位与规划：模糊语义导航智能体将公园名称转化为具体坐标，调用地图API生成导航路线。
- ⑤ 多模态反馈：系统通过TTS（可能由MiMo-V2-TTS或ZipVoice驱动）用拟人化语音回复，并在车机屏幕和HUD上显示路线。

MiMo端侧大模型/MiMo-V2-Omni（全模态基座模型）

- 多模态理解能力升级主要体现在以下方面
- 可视即可问：超级小爱从识别车辆、建筑到识别万物，多模态识别能力覆盖超多场景。无论是路边的动植物、交通标识、还是眼前的自然景观，小爱均可实时识别并解答疑问；小爱同学能识别路标、车辆、动植物、云朵等，这需要强大的多模态感知能力。
 - 语音跨端互控：小爱跨平台、跨设备执行复杂任务，实现车机与手机，例如，“小爱同学，在手机美团里给三里屯的海底捞打个电话”，最终落地执行，很可能通过名为 Xiaomi midaw 的端侧AI智能体来实现



- 小米XLA智能大模型是小米为智能驾驶（HAD）打造的专属底层架构，核心服务于智驾。其中X 多模态认知输入主要包括雷达、麦克风、摄像头、激光雷达、GPS、互联网信息、机器人数据，主要作用是给智驾提供全感官输入，比如听到后方鸣笛自动取消变道、结合路况 / 地图做全局决策；L 高效的原生语言基于小米的具身基座模型（MiMo-Embodied），用大模型做驾驶决策推理，而非机械执行规则；A 拟人化的行为用世界模型做虚拟仿真、强化学习优化驾驶策略，让车开得更像老司机。
- 小米XLA智能大模型已量产上车2026年推出的新一代小米SU7

➢ 老款 SU7 的语音以座舱功能控制为主（如空调、车窗、音乐），无法直接指挥驾驶动作；2026 款新增语音控车，可通过自然语言控制变道、跟车距离、超车、指定车位泊车等驾驶动作，且支持车外语音泊出、边走边说、连续多意图

整个流程：用户发出指令“变道（右）”→ASR→语义解析→意图“变道（右）”→发给 XLA→XLA做出驾驶决策→车辆执行

- 超级小爱：负责听、听懂、确认用户的意图，然后把明确指令传给智驾系统；
- 智驾XLA 大模型：拿到指令后，结合实时路况、传感器数据做安全判断和驾驶决策；
- 最后车辆执行机构才真正去变道、跟车、泊车



小米miclaw有望成为下一个跨场景、跨终端的系统级AI智能体

- Xiaomi Miclaw (龙虾) 是小米于 2026 年 3 月 6 日 推出、基于自研 MiMo 大模型打造的系统级 AI 智能体 (Agent) ，也是国内首个对标 OpenClaw 架构的手机端 AI 交互产品目前处于手机小范围邀请制封测阶段，卢伟冰表态，miclaw 未来将拓展到电脑、手表端，最终服务于小米"人车家全生态"设备
- MiClaw被小米高层定义为未来操作系统的雏形，核心在于它推动了AI从 ‘对话能力’ 向 ‘系统级执行能力’ 的实质性跨越，并深度融入小米的 ‘人车家全生态’ 。

Xiaomi Miclaw

系统级AI智能体

让手机成为你的智能助手，基于小米MiMo大模型，实现自主拆解任务、调用50+系统工具与米家生态设备，完成复杂执行

粉丝站点，并非官方站点

- 正在内测中
- 了解同类PC产品Molili

区别传统语音助手（仅能简单指令），Miclaw 能 "思考、规划、执行、反馈"，自主完成复杂任务

01

系统级操作能力

- miclaw 打通了超过 50 种系统级工具接口，它可以读取短信、日历、通讯录、通话记录；操作相册、管理文件；创建日程、设置闹钟；发送消息、拨打电话；获取位置、查询天气

02

个人上下文理解

- miclaw 有一个记忆系统，它会记住用户的习惯偏好（几点起床、喜欢什么口味）；人际关系（家人、朋友、同事）；历史决策（上次选了哪家餐厅）；
- 它能理解"老地方见"是什么意思，也能在你问"明天怎么安排"时，把日历、天气、待办串在一起给你答案。

03

动态接入第三方应用，无需 APP 适配

- Miclaw 基于 MCP （Model Context Protocol）开放协议，动态接入第三方应用，无需 APP 适配。系统级权限直接调用功能，非模拟点击，安全稳定、不被判定外挂。

04

自进化能力

- Miclaw 记录用户习惯、偏好、使用场景，自动优化任务执行。常用指令记忆、偏好设置、场景适配，长期使用更贴心、更高效。多设备协同，控制米家生态、手机、电脑、汽车，全场景互联

以系统级 AI 智能体实现智能跨端调度，构建统一人车家全生态中枢

佐思汽研，致力于汽车新四化的行业研究、 咨询规划、数据监测和公众号推广



zuosiqiche