

Agent时代

鸿蒙应用生而智能

——鸿蒙智能体框架白皮书



Agent

2025年6月



邬贺铨

鸿蒙智能体：自主创新引领智能生态新纪元

当前，全球数字技术正经历智能化、泛在化的深刻变革，人工智能与万物互联的深度融合，推动着产业生态的颠覆性重构。在这一进程中，鸿蒙智能体的崛起，不仅是中国信息产业自主创新的重大突破，更是全球智能终端生态迈向新范式的重要里程碑。作为新一代操作系统与 AI 技术的集大成者，鸿蒙智能体以技术自主为根基、生态协同为路径，为中国数字经济发展注入新动能，也为全球产业智能化升级贡献了中国智慧。

鸿蒙智能体的突破性在于其开创了“以人为中心”的操作系统新范式。通过分布式架构与原子化服务能力，实现跨终端、跨场景的智能协同。例如，在智能家居场景中，用户无需手动切换设备，鸿蒙可通过意图感知自动联动灯光、空调、安防等设备，形成“服务随人动”的无缝体验；在工业领域，其微内核设计和高可靠性特性，为智能制造、智慧能源等场景提供了毫秒级响应的底层支撑。这种“软硬件解耦、服务原子化”的技术路径，使得终端设备从“功能容器”进化为“智慧体”，为 AI 应用的场景化落地开辟了全新空间。

鸿蒙智能体的核心价值，在于构建了“开放共赢”的生态体系。通过开源项目 OpenHarmony，华为联合超 5000 家合作伙伴，打通了从芯片、模组、终端到应用的全产业链条，形成了全球第三大智能终端生态圈，其“一次开发、多端部署”的开发模式，更将应用开发效率提升 60% 以上，让中小开发者也能快速融入 AI 创新浪潮。在智能汽车领域，鸿蒙车机操作系统已赋能多家车企，推动中国智能网联汽车技术领先全球。这种“技术普惠 + 生态共建”的模式，为实体经济数字化转型提供了可复制的中国方案。

鸿蒙智能体的发展承载着国家数字主权的战略使命。技术层面，鸿蒙智能体框架 HMAF 在传统软件安全工程的基础上，迭加构建 AI Agent 典型安全防线，实现全栈安全自主可控；产业层面，鸿蒙生态实现了从底层芯片到上层应用的完整技术栈，使我国首次在基础软件领域摆脱了对西方体系的依赖。更为深远的是，鸿蒙通过定义智能终端交互标准，在消费电子、智能汽车、工业互联网等关键领域推动“中国标准”走向世界。这种“技术自主 + 标准引领”的双重突破，与“数字中国”战略形成深度共振，成为提升国际竞争话语权的关键支点。鸿蒙智能体的成功实践，标志着中国的终端操作系统从“技术跟跑”到“生态领跑”的历史跨越，它不仅证明了自主创新道路的可行性，更展现了开放协同生态的旺盛生命力。

我们期待，鸿蒙智能体生态能够持续深化产学研用融合，为中国式现代化筑牢数字基座。这既是时代赋予中国科技界的重任，也是全球智能化浪潮中不可或缺的东方力量。鸿蒙之路，基于使命，始于技术，成于生态——为中国乃至世界打开万物智联的无限可能。

邬贺铨



余承东

华为常务董事
终端 BG 董事长

随着人工智能技术持续突破，特别是大模型、全场景融合感知与端云协同等能力的发展，智能终端正经历从“工具”向“协作者”的深刻变革。这场变革不仅改变了人机交互的方式，更重构了整个智能生态的服务逻辑。消费者的需求不再是简单的功能堆叠，而是追求随时随地、自然流畅、无感协同的智慧体验。他们希望终端设备更聪明、更懂人，不仅能理解指令、完成任务，还能超前思考、主动服务。在这样的趋势下，智能体成为新时代智慧体验的关键载体，以用户意图为中心，以多模态交互、意图理解、任务闭环与持续学习进化等关键能力为特征，打通用户、系统与服务之间的链路，带来全新的服务模式。

华为终端致力于打造王者产品，构筑智慧全场景极致体验。我们持续通过系统架构与软硬芯云垂直整合创新，将 AI 融入到鸿蒙操作系统底座，让 AI 成为驱动新一代智慧体验的关键引擎。去年华为正式发布全场景智能操作系统 HarmonyOS NEXT，率先将 AI 与操作系统深度融合，小艺也正式升级为“系统级智能体”，无处不在、随时可被召唤，从而将 AI 的便捷性和实用性提升到了一个新的高度。现在，我们进一步提出鸿蒙智能体的技术和生态发展主张。鸿蒙智能体是对传统智能体的升级，实现了系统级安全可信、多智能体高效协作、自主可控的个性化、多设备间自由流转等独特能力，引领系统与应用交互协同的新范式。

鸿蒙应用生态正蓬勃发展，为赋能鸿蒙应用生态智能化演进，我们提出鸿蒙智能体框架，以多样化的交互入口、高效可信的通信协议、丰富的开发工具生态、全链路的安全防护为基石，构建一个开放、灵活、可演进的智能体 + 鸿蒙应用开发平台。我们坚信，鸿蒙智能体框架将加快生态伙伴的创新脚步，不论是在手机、PC、车机、智慧屏、穿戴设备，还是在办公、家庭、出行等不同场景中，鸿蒙应用智能体都能实现个性化、主动化、可信赖、一致的智慧体验，拉近服务与用户的距离，持续推动鸿蒙智能的体验变革。

鸿蒙智能体时代已经开启，我们期待与广大鸿蒙应用开发者和合作伙伴一道，持续推动鸿蒙智能体技术创新与生态共建，为用户带来“服务合时宜、服务不间断”的智慧新体验，共创鸿蒙万物智联的未来新世界。



鸿蒙应用智能的愿景

近年来，生成式人工智能（Generative AI）迎来前所未有的技术飞跃，以大语言模型（LLM）为代表的基础模型进入高速迭代期，一路狂奔，不断刷新能力边界。模型在复杂任务推理、规划、代码生成等方面取得显著提升，在各类权威基准测试中不断突破记录，多项能力指标接近甚至超越人类专家水平。

模型的上下文窗口显著扩展，主流模型已支持十万到百万级 Token 输入，可处理长文档、复杂多轮对话，实现“全局理解”和“文档级记忆”，为报告审阅、代码仓分析等高复杂度任务提供了技术基础。

与此同时，具备跨模态能力的理解与生成模型加速涌现，支持文本、图像、语音、视频等多模态输入输出，大幅提升了对现实世界的感知与表达能力，赋能教育、创意设计、广告营销、媒体内容生成等新场景，成为新一代人机交互与智能创作的核心。

知识蒸馏、低比特量化等技术推进了模型压缩和高效推理，使大模型“端侧可部署”成为现实。通过“本地小模型 + 云端大模型”的协同推理，可以最大化发挥“端侧快”和“云侧强”的优势，同时解决信息安全隐患、云端算力成本过高等问题。

这一系列技术的突破，正在重塑人类与数字世界的交互方式。生成式 AI 不再是问答工具，而是嵌入到每一个终端、系统、应用与服务中的可持续演化的智能体（Agent），具备感知周围环境、理解用户意图、学习用户偏好、自主执行复杂任务的全流程能力。

2024 年 6 月，华为向业界发布《AI 终端白皮书》^[1]，明确指出智能体将引领终端进入全新人机交互及智能协作的时代，如多模态大模型带来更自然、更全面的人机交互体验，AI 自主化程度越来越高，“以意图为中心”的 AI 与人协作方式正在建立。与此同时，消费者对于终端设备的使用习惯也在悄然改变，需要更加智能化、互联化和人性化的全场景智能服务体验。未来，多设备、多场景的使用将成为消费者的必要选择，无论是工作、生活、学习还是娱乐，都能在不同的终端设备间无缝切换，享受连贯、高效的服务。

生成式 AI 加速与终端产品深度融合，不仅仅为消费者带来了全新的人机交互体验，也推动了应用形态和生态范式的变革。华为在《鸿蒙 2030 白皮书》^[2] 中提出，当前主流操作系统的生态范式均以 APP 为中心构建，未来几年的鸿蒙生态系统将以面向场景的智能服务形态为终极演进目标，呈现出以 AI 智能体为中心、多种服务形态并存的新范式，包括 APP、元服务等已存

在的形态以及以 AI 智能体为载体的智能服务新形态。各种形态之间相互协同、相互融合，为用户提供更优的体验。其中，AI 智能体可以根据意图分解的任务目标，自主调用相应的 APP、元服务等以高效完成目标，同时已有的 APP、元服务也可通过智能化升级获得 AI 智能体的能力，提供更智能的服务体验。进一步，系统智能体将与各种应用智能体通过智能体联邦的形式互相协作，共同完成更复杂的任务，而无需用户再跨多个 APP 和元服务进行操作。

在新的生态范式下，智能体将以惊人的速度渗透进各行各业，推动一场新的生产力与创造力革命，成为价值创造者，深刻影响着人们日常生活、工作、学习、娱乐的方方面面。面对这一变革浪潮，行业亟需一个有机协调、多方共赢的系统性框架，以释放智能体之间的协同潜力，基于此，在智能体时代，鸿蒙核心愿景是构建一个全新的智能体价值网络（Agent Value Network），在这个网络中，所有智能体将协同工作、共同创造并共享收益。鸿蒙系统将持续构筑统一开放的智能体 OS 底座能力，全面赋能鸿蒙应用智能化演进，为每一款应用注入更自然的多模态交互能力、更高效的跨应用协同体验以及更深层的系统级感知和触达能力，助力每一款应用“生而智能”。同时，鸿蒙系统智能体小艺作为用户交互入口，将构筑精准的用户意图识别能力和公平透明的流量分发机制，让每一个参与协作的鸿蒙智能体或智能化鸿蒙应用，都成为价值创造者，让每一次协同都产生增量收益。

01 行业趋势：应用智能化和 Agent 产业趋势 01

- 1.1 从被动响应转向主动服务，智能体驱动人机交互范式变革 02
- 1.2 智能体爆发式增长，系统智能体和垂域应用智能体双轨发展，走向协同 04
- 1.3 一站式开发框架和技术标准化加速智能体构建和商用落地 06

02 鸿蒙智能体的定义 08

03 鸿蒙智能体框架（HMAF） 11

- 3.1 智能体全新交互 13
- 3.2 智能体协议升级 18
- 3.3 智能体高效开发 24
- 3.4 智能体安全可信 28

04 总结与未来展望 31

05 参考资料 33





01

行业趋势：
应用智能化和 Agent 产业趋势

1.1 从被动响应转向主动服务，智能体驱动人机交互范式变革

1.1.1 AI 从“被动响应”的工具调用转向“主动服务”的成果交付

随着生成式 AI 技术的飞速发展和规模化使用，AI 系统不仅具备了强大的多模态理解、推理与生成能力，更内嵌了工具调用、长期记忆、代码执行等机制，能在无需人工干预的情况下感知环境、规划行动并完成复杂任务，实现从感知到决策再到执行的闭环，从被动响应式工具演进为可自主执行复杂任务的 AI 智能体。例如，以华为小艺智能体、Manus、智谱 AI AutoGLM、OpenAI Operator 等为代表的智能体，标志着 AI 开始从“对话交互”向“任务闭环”进化，未来将会有更多智能体介入现实世界的操作。

华为在《AI 终端白皮书》中提出，当 AI 结合了更高维度的感知结果，将触发思考，进而影响自主行为。与人类常见的行为模式一样，看到乌云即可联想到下雨从而带了一把伞出门，AI 将学会感知复杂世界并与之互动，并在此过程中触发智能的、可按需编排调度的主动服务。在不久的将来，更高维度的智能感知、更智慧的基础模型、更高阶的空间智能将为具身智能赋能，开启智能社会的更多可能性。

为了让消费者对终端的 AI 能力有更清晰、更直观的认知，同时也为了让产业界对 AI 终端的能力演进达成统一的共识，牵引终端系统智能化的演进，基于人与 AI 在任务中的协作关系、AI 的参与程度，并参照汽车驾驶自动化分级的理念，以及清华大学 PERSONAL LLM AGENTS^[3]（个人大语言模型智能体）中的智能体能力分级策略，我们提出了终端智能化体验的 L1~L5 分级标准。当前，L3 可自主闭环任务的 AI 智能体正在加速发展，预计 2030 年终端系统有望实现从低阶到高阶的智能化跃迁，达到 L4+。

AI 智能化分级		AI	人
L5	智慧级-自主智慧	AI 超越人类专家水平的能力，全面自主	人类授权
L4	指导级-专业指导	AI 提供达到人类专家水平的定制化服务	人类参与
L3	协作级-协作自治	AI 自主拆解及分配任务，闭环执行	人与 AI 协作，人类监督
L2	任务级-任务执行	AI 执行被分解的任务	人类拆解及分配任务
L1	功能级-辅助工具	AI 作为工具被调用	人类执行并闭环任务

图 1：终端智能化分级标准

Gartner 也在《2025 年十大战略技术趋势》^[4]中指出，目标导向驱动着 AI 系统的自主性和适应性更强，完成任务的范围和复杂性将不断扩大。当前 AI 助手和大语言模型（LLM）虽能执行文本生成、内容摘要或基础工具使用等任务，但无法基于“自主意识”采取行动。自主智能体正在改变这一现状，现有及新兴 AI 技术（如智能体框架、多智能体协作、AI 安全护栏和工具调用功能）正赋予基于 LLM 的系统更多的自主性，即在最少的人工监督下、适应环境并在复杂环境中自主执行任务并实现目标，这将极大拓展 AI 的应用潜力。例如，Agentic AI 可分析数据、开展调研、编制任务清单，继而通过 API 或机器人系统在数字 / 物理世界执行这些行动。Gartner 预测，到 2028 年，至少 15% 的日常工作决策将由代理型 AI 自主做出，而 2024 年这一比例为 0%。

1.1.2 智能体驱动“以意图为中心”的人机交互范式变革

随着 AI 智能体自主能力的持续进化，人机交互范式将发生变革。从传统的“以用户指令为中心”发展为“以用户意图为中心”，用户只需要表达出需求或期望，而将实现需求的过程交由智能体系统来完成，为人类带来更深层次的便利和效率提升^[1]。

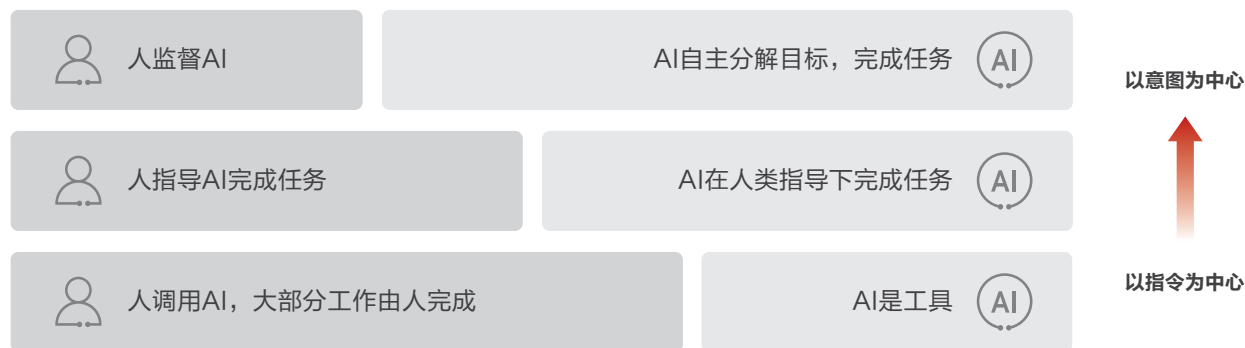


图 2：“以意图为中心”的人机交互范式

- ◎ 互联网时代：人机交互主要采用“用户输入 - 系统处理 - 结果输出”的线性交互模式，用户通过点击、滑动、输入等物理操作与应用 UI 界面互动，本质上是一种“指令驱动”的被动服务，人机交互是一对一顺序进行的指令性交互。在这个过程中，由用户完成绝大部分操作，AI 仅作为工具进行信息处理和建议。

- ◎ 智能体时代：AI 将通过多模态融合感知（语音、文本、视觉及传感器信息等）实时捕捉用户场景化需求，并深度解析潜在意图，实现“意图驱动”的主动服务。例如：智能体可根据用户生活习惯，在适当时机主动推送个性化建议，如在临近晚餐时间提醒用户用餐，并根据用户的口味提供给用户合适口味的晚餐的选择；甚至自动整合用户多个应用信息，为用户规划出行行程，如预订车票、机票和酒店，并根据当地天气与用户偏好推荐旅游景点和特色美食，一站式满足用户的出行需求。

人机交互范式，从功能机时代的物理按键，到智能机时代的触屏操作，再演进至 AI 智能体时代的意图交互，这一历程体现了“从形式操控到语义理解”的技术进化，其本质是机器对人类认知和行为习惯的持续适配。

在“以意图为中心”的交互范式下，用户角色从“功能操作者”转变为“目标委托者”，智能体则从“工具”升级为能独立完成任务并交付结果的“数字协作者”，移动终端设备作为用户与智能体交互最直接的入口，其系统交互入口将演进为承接、理解和分发用户意图的关键节点，本质是触发万千智能体协同工作的调度系统中枢。在这一架构中，系统智能体承担着“中控大脑”的角色，需对用户意图具备极高的理解精度，能够将复杂任务合理拆解，并精准调度下游的应用智能体协同完成各类子任务。根据华为实验室 2025 年 5 月评测数据，鸿蒙系统智能体小艺在复杂任务规划与执行场景，基于用户意图理解的工具调用和智能体分发的准确率在单轮对话场景已超过 95%，多轮对话场景已超过 90%，在意图不明场景的用户澄清准确率也已超过 90%。未来，小艺将继续深耕，探索和构筑更加精准的系统级用户意图融合感知、识别和分发调度能力。

1.2

智能体爆发式增长， 系统智能体和垂域应用智能体双轨发展，走向协同

DeepSeek 等大模型技术的突破性进展，使 AI 智能体具备了更强的理解、规划和执行能力，应用场景更加广泛。行业普遍认为 2025 年将成为 AI 智能体商用爆发的“元年”，市场规模呈现爆发式增长趋势。市场研究机构 Grand View Research 数据显示^[5]，2024 年全球 AI Agent 市场规模达 54 亿美元，预计 2025 - 2030 年将以 45.8% 的年复合增长率增长，到 2030 年达到 503.1 亿美元。

从整体应用格局来看，AI 智能体形成 " 系统智能体 + 垂域应用智能体 " 的双轨发展模式。

系统智能体成为用户流量新的超级入口已成定局。华为在《鸿蒙 2023 白皮书》中提出，智能设备的交互入口将可能从基于应用程序的界面逐步演变为基于智能体的界面，向用户提供个性化场景化的智能服务。系统智能体成为既智慧又懂你的虚拟 AI 私人助理，具有记忆和自主学习能力，随着陪伴用户时间的增长，对用户的过去历史、当前状态、个人喜好等越来越了解，可以根据用户的意图代表用户处理生活和工作上的日常杂务，自主完成各种任务。例如，用户想计划一次旅行，AI 私人助理根据用户的喜好和预算，做出深度个性化旅行建议方案，并自动预定用户喜爱的酒店、航班等。2024 年被认为是生成式 AI 手机的“元年”，作为行业先驱者，2024 年 10 月 22 日华为正式发布全场景智能操作系统 HarmonyOS NEXT，率先将 AI 与操作系统深度融合，小艺也正式升级为“系统级小艺智能体”，无处不在、随时响应，从而将 AI 的便捷性和实用性提升到了一个新的高度。2025 年 3 月 20 日华为 Pura X 手机发布，全新小艺重磅升级^[6]，兼具情感与效率，拥有更具呼吸感的视效与更拟人的真实感音色，懂用户的情绪与生活，同时能够自适应开启深度思考，用更强大脑帮助用户解决更多复杂的问题。据统计，2025 年 5 月鸿蒙系统智能体小艺的使用率已超过 30%，可见，小艺越来越受到用户的喜爱，这一数据也将随着小艺的进化进一步提升。

垂域应用智能体快速发展，将成为下一个 C 端应用主战场。垂域应用依托行业数据壁垒，通过 AI+ 行业解决方案端到端解决用户痛点，并形成私有数据飞轮，易构筑更坚实的护城河。据德勤分析，当前，AI+ 教育类应用已率先普及，AI+ 法律、AI+ 财经、AI+ 医疗等解决方案类应用也存在较大增长潜力。初创企业通常优先切入生产力场景，以内容创作、垂直专业、工作效率类 AI 应用抢抓用户流量。其中，AI 图像 / 视频处理、办公助手等工作场景 AI 应用的降本增效价值受用户广泛认可，设计、剪辑、内容制作等用户群体内已实现较高渗透。

用户总是习惯于在同一个入口实现多种需求，而每一个独立智能体的能力都是有限的。未来，移动终端设备的系统级交互入口将承接用户的各种复杂意图，成为统一的多智能体服务入口。系统智能体与众多垂直领域应用智能体将走向深度协同，通过相互协作共同完成更多场景下的更复杂的用户需求，形成全新的智能体价值网络。在这个网络中，每一个应用智能体都是价值创造者，也是利益分享者，其核心竞争力体现在是否能够帮助用户解决特定场景下的实际问题，而系统智能体需要构筑精准的用户意图识别能力和公平透明的流量分发机制，让每一个应用智能体都能在协作中创造价值、分享收益。构建统一开放的智能体通信协议，是打造智能体价值网络的前提。

这类似于，移动终端设备不再仅仅是提供一个“操作系统”让开发者在上面运行程序，而是构建了一个“数字经济体”（智能体网络），让不同的“企业”（Agent）在这里遵循统一的“市

场规则”（协议），进行“资源交易”（能力交换）和“产业链协作”（任务协同），最终实现整个经济体利益的最大化。

1.3 一站式开发框架和技术标准化加速智能体构建和商用落地

随着智能体迎来爆发式增长，智能体开发框架 / 平台也大量涌现，包括华为、字节、百度、阿里、腾讯、讯飞等企业，纷纷推出自己的智能体开发框架或平台，提供可灵活调用的模型 / 工具以及一站式开发、调测能力，大大降低了智能体的开发门槛。

然而，智能体从原型开发到商用落地，仍然面临着一些挑战。如何打造高可靠的智能体，华为和业界的实践指出了一些关键方向：

- ◎ **通过记忆让智能体“学会持续进化”**：一方面，记忆可以让智能体在多轮交互中记住用户历史意图，避免信息丢失，保证语义连续性，提高长期任务执行稳定性；另一方面，长期记忆可用于存储任务经验、领域知识、用户偏好，提供更可靠更个性化的服务。例如，小艺通过建立一个存储并管理用户信息的记忆体，可以在解答问题、执行操作时进一步结合用户习惯，给出更具个性化的回答和反馈
- ◎ **通过工具让智能体“能做真实工作”**：通过接入更多的业务数据源、调用更多的垂域工具 / 服务，可以扩展智能体的能力边界，提升搜索、数据查询、计算分析等能力，完成更多场景更高复杂度的任务。例如，鸿蒙操作系统通过构建统一开放的意图框架，实现了系统级的跨应用跨服务的数据和能力共享。业界也推出 MCP (Model Context Protocol) 协议^[7]，为大模型提供了标准化的数据源和工具接入方式，大大降低了工具集成和升级维护成本，促进了智能体工具生态的繁荣。
- ◎ **通过多智能体协同让智能体“像团队一样稳健执行复杂任务”**：通过结构化分工，让每个智能体执行擅长任务，多智能体间互补协作，可以更好的发挥群体智能，更稳健地执行复杂任务。华为提出小艺系统智能体和垂域应用智能体协同方案，帮助用户完成更多复杂任务，本质上就是对多智能体协同的一个最佳实践。同时，华为提出鸿蒙 Agent 通信协议，定义了鸿蒙系统下多智能体之间的交互规范，以实现高效、安全协作。业界如谷歌也推出 A2A (Agent to Agent) 协议^[8]，旨在让不同平台的智能体可以通过基于任务的交互来完成协作。

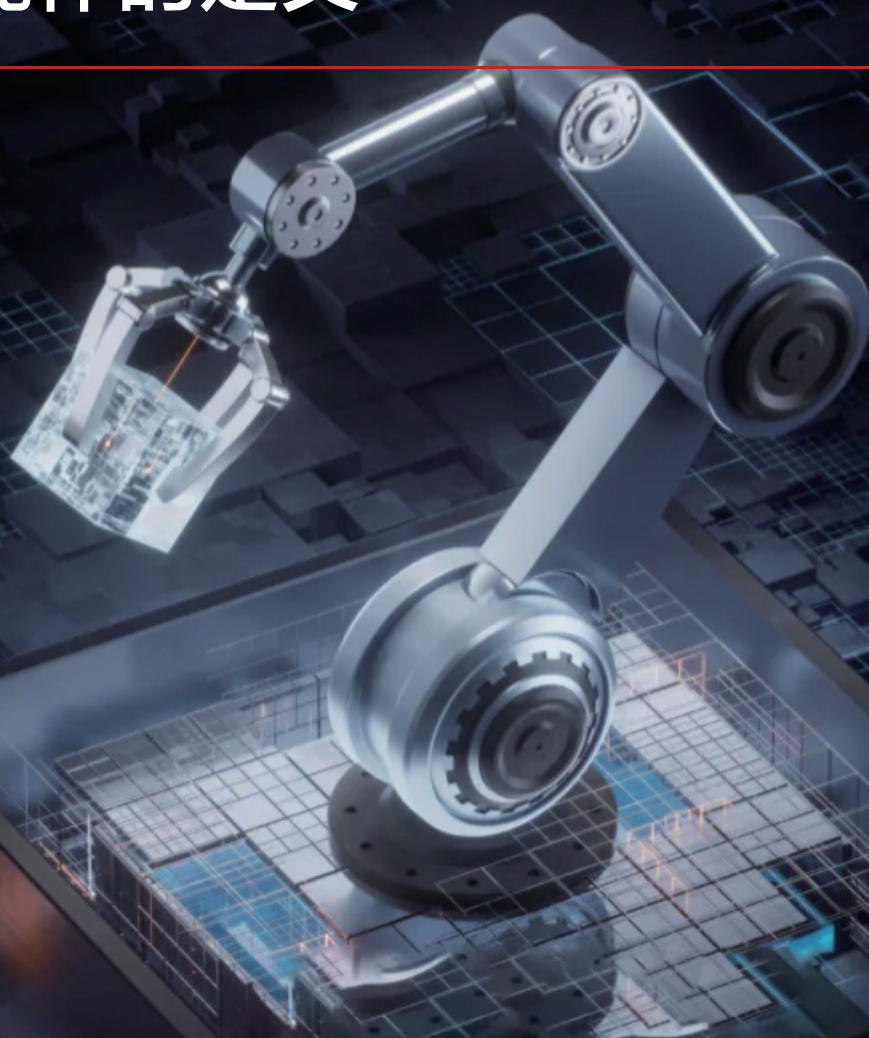
标准化的工具协议和智能体通信协议为智能体的能力拓展提供了坚实基础。鸿蒙意图框架、MCP 等工具协议实现了垂域工具和服务的标准化接入，显著降低了智能体与外部系统的集成复杂度；而智能体通信协议则为多智能体间的任务协同建立了通用语言，使得跨平台、跨系统的智能体协作成为可能。这些技术标准的建立和不断完善，不仅加速了智能体生态的互联互通，更推动了开发资源向功能创新而非兼容适配的集中投入。

智能体开发框架的持续完善，推动智能体开发进入规模化生产阶段。开发者依托一站式平台可以快速实现记忆集成、工具调用和多智能体协同等核心功能，并通过可视化调测工具完成智能体端到端效果优化。这种标准化、模块化的开发模式，将有效降低创新门槛，推动各类垂域应用智能体的高效构建和规模化商用。

移动终端设备厂商能够为智能体开发提供更多独特的系统级能力。例如，通过跨设备跨应用信息的全局记忆，为智能体提供更加全面的上下文信息，实现更精准的用户意图识别和应用智能体协同；通过开放的意图框架接入并共享丰富的端侧应用工具，使系统智能体和应用智能体能够实现更无缝的跨应用操作体验；同时，统一账号体系以及端到端的安全与隐私保护框架，进一步保障多智能体的安全高效协作。这些差异化能力将成为终端厂商在智能体时代赋能开发者、构建竞争力的关键。

02

鸿蒙智能体的定义



现有基于大模型的智能体具备“可以自主地理解意图、规划决策、执行任务、调用工具，并具有记忆能力”等特点，在此基础上，鸿蒙智能体增强了与鸿蒙操作系统、应用的协同，可以“与系统、应用和元服务无缝协作实现复杂任务，形成智能体价值网络”。概念上，鸿蒙智能体分为两类：

- ◎ **系统智能体**：小艺作为鸿蒙系统中的系统级智能体，可以充分结合操作系统底层能力，为用户提供体系化、可扩展的智能能力；通过系统级的感知能力、记忆能力和工具能力，小艺系统智能体像一位私人助理，一直关心、实时在线、适时服务，站在距离用户最近的位置，协同其他智能体为用户闭环复杂任务，代表用户超前思考，为用户带来智能的服务体验。
- ◎ **领域智能体**：类似于终端操作系统（HarmonyOS、Android、iOS）的系统应用和不同领域应用一样，除了系统智能体之外，下一代 AI 终端中还将存在多个领域智能体。领域智能体具有特定领域的专属知识和工具，更擅长闭环垂域任务。系统智能体和领域智能体相互协作，为用户提供综合性与专业性结合的服务体验。由于当前的鸿蒙应用开发者和企业具有各领域丰富的积累，我们判断大量的高质量领域智能体将首先由应用厂商孵化或者转型而来。这种与鸿蒙应用\元服务存在派生与协同的领域智能体，我们统称为“应用智能体”，是鸿蒙生态的重要组成部分。

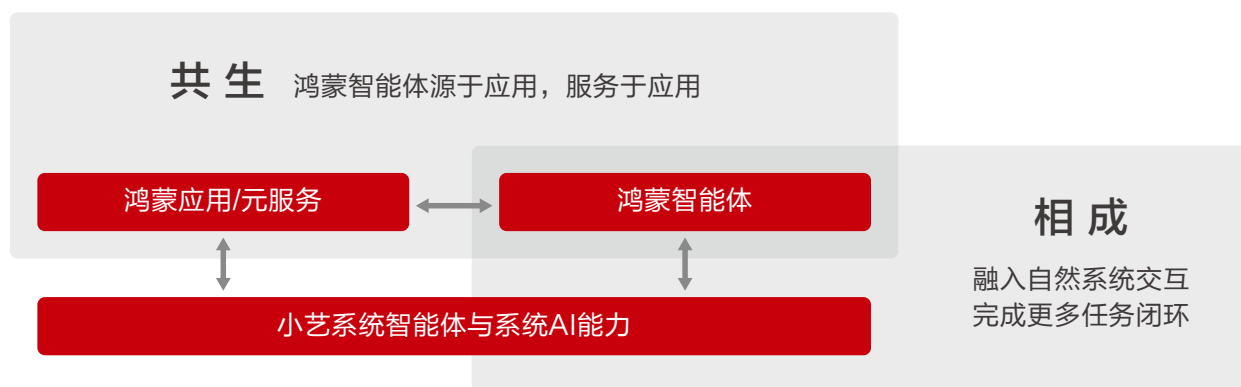


图 3：鸿蒙智能体与鸿蒙系统、鸿蒙应用的关系

基于鸿蒙智能体，消费者与鸿蒙系统、应用的交互将发生根本改变，通过鸿蒙智能体系统级安全可信、自主可控的个性化、多智能体间高效协、多设备间自然流程等特质，真正从传统的“以用户指令为中心”转变为“以用户意图为中心”：

- ◎ **系统级的安全可靠：**随着智能体的广泛应用，以及其规划决策和执行任务的特点，智能体时代面临远超过去移动应用时代的安全可信挑战。为此，鸿蒙提供了全栈的安全防护与系统级的安全可靠，具体可参考 3.4 节。
- ◎ **多智能体高效协作：**随着技术的进步与用户的适应，交给智能体的任务会越来越复杂，愈发需要通过多智能体协作让智能体“像团队一样稳健执行复杂任务”。为此，我们发布了鸿蒙 Agent 通信协议并升级了意图框架，以实现多智能体之间的高效协作，具体可参考 3.2 节。
- ◎ **自主可控的个性化：**无论是系统智能体与应用智能体之间的协同，还是多智能体间的高效协作，必然伴随着用户数据的传输（如系统智能体对用户意图与使用偏好有更精准的理解，需要基于此与领域智能体沟通）；为此，我们提出了自主可控的个性化，基于“数据最小化”、“用户可知可控”等原则在智能体之间传递数据。
- ◎ **多设备间自由流转：**在鸿蒙智能体设计之初，我们就目标生态智能体可以做到一次接入，在华为终端全场景 1+8 设备上上进行多设备分发；同时不同设备上的同一智能体也能充分协同工作，共享上下文。

我们认为，未来服务分发模式会发生变革，从人找服务，到服务找人；从离散性服务，到连续性服务；从被动调用、到主动服务。而鸿蒙智能体正是支撑这一变革的核心要素。

对消费者的价值：过去通过人进行 APP 操作串联，记不住复杂的操作步骤和不常用应用；如今可以通过自然语言交互，系统智能体与应用智能体进行协同完成更复杂任务。在不同场景、不同设备上用户可以更便捷地获取服务，享受场景化智慧化的“服务合时宜、服务不间断”的新体验。

对生态伙伴的价值：生态伙伴的服务分发入口将获得极大丰富，可直接被系统调用并完成闭环，服务的触达路径更短更直接，商业转化效率更高。同时基于大模型能力和智能体平台，开发者可以将复杂的业务逻辑通过简单的方式内化到智能体的调用流程里，大大降低开发门槛，让鸿蒙开发更快更简单。



03

鸿蒙智能体框架（HMAF）

在未来应用与智能体走向深度融合与协同的大趋势下，我们提出面向鸿蒙生态的**鸿蒙智能体框架（Harmony Agent Framework, HMAF）**。该框架全方面定义了鸿蒙系统下操作系统、鸿蒙应用\元服务与智能体的交互协同范式，致力于构建一个具备自主决策和群体协作能力的 AI 生态体系，赋能三方生态应用智能化升级，为各领域的应用和智能体开发者提供架构和框架支撑。框架核心包含以下四个方面：

- ◎ **智能体全新交互：**在鸿蒙系统下，智能体将和系统交互深度融合，小艺系统智能体、系统级分发入口将与智能体充分互通，基于用户意图快速直达相关的智能体。智能体可以通过语音、卡片、应用伴随的 OnApp 和 InApp 组件等方式与用户自然交互。
- ◎ **智能体协议升级：**提出鸿蒙 Agent 通信协议，实现鸿蒙系统内的智能体之间高效通信、多智能体协同。对鸿蒙意图框架全面升级，在兼容 MCP 协议的基础上，增强智能体与鸿蒙应用、智能体与鸿蒙系统的工具、插件能力对接。最后，针对应用开发者提供鸿蒙 Agent Kit，允许鸿蒙应用与智能体深度集成，实现智能体与应用伴生增强。
- ◎ **智能体高效开发：**通过小艺开放平台全面开放鸿蒙系统 AI 和插件能力，赋能智能体开发。将 AI 辅助智能体开发融合到智能体开发流程，提升智能体开发效率。支持智能体开发与鸿蒙应用/元服务开发 IDE 工具协同，完成鸿蒙智能体与应用的高效协同开发。支持鸿蒙智能体与华为全场景终端设备真机联调，便捷发布上架至智能体市场，提供鸿蒙生态体系下的智能体全生命周期版本管理。
- ◎ **智能体安全可靠：**鸿蒙系统提供面向智能体的全栈安全防护，通过可信基础设施的硬件安全、智能体数据安全、智能体端云执行环境安全、智能体统一身份认证安全等防线，构建鸿蒙智能体框架下的全链路安全体系。

鸿蒙智能体框架（Harmony Agent Framework）：Agent时代，应用智能新范式

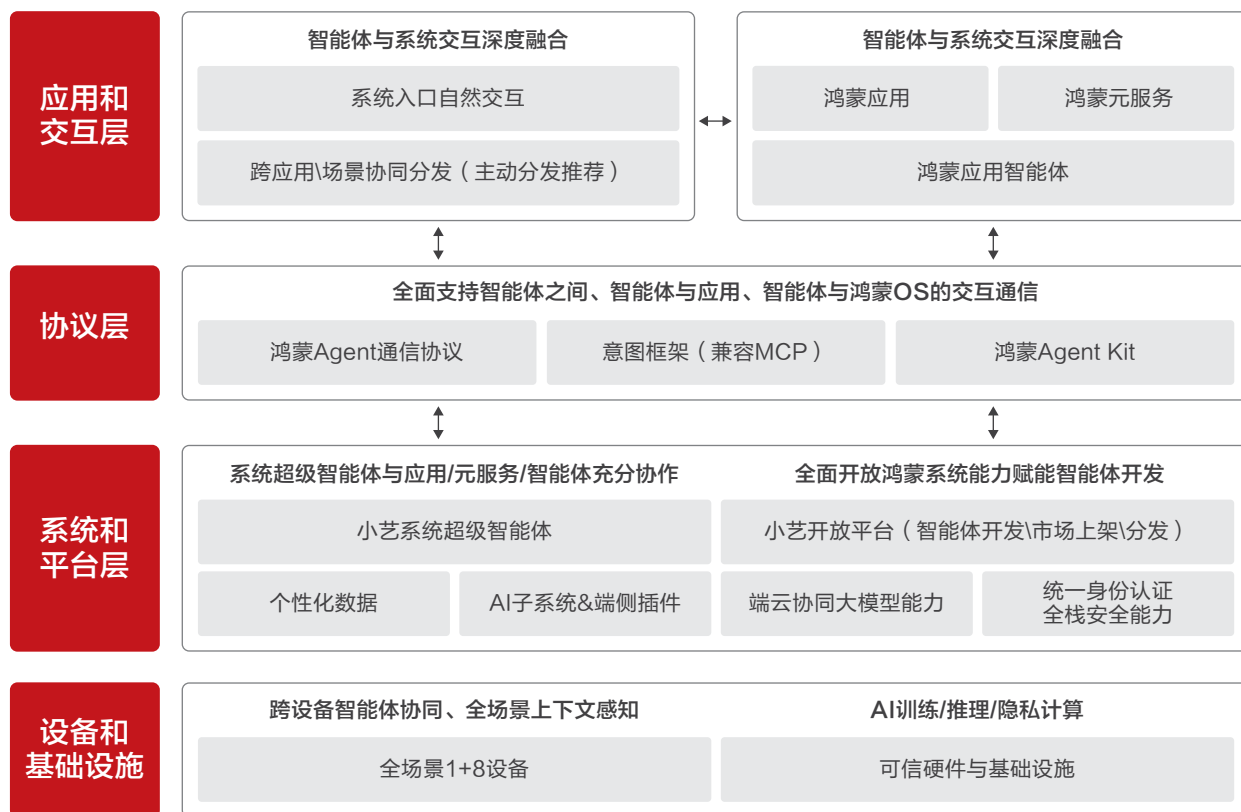


图 4：鸿蒙智能体框架 HMAF 介绍



3.1

智能体全新交互

伴随着智能体生态的蓬勃发展，智能体的数量越来越多，用户触达智能体成本逐步升高，智能体功能也越来越丰富，却没有带来用户智慧体验感的升级，亟需更加丰富的便利的智能体分发方式和全新的交互范式。

当前应用智能体的核心体验缺陷：

- 1) 入口深：智能体入口深度耦合于应用内层级结构，依托于用户进入应用后，主动找到智能体功能。无法便捷快速触达。
- 2) 范式缺失：用户认知成本高，交互框架未标准化，除搜索型智能体外，缺乏普适性强的基础交互范式。用户需重复学习不同的智能体操作逻辑。
- 3) 跨应用服务断裂：单个智能体受限于单个应用，无法构建端到端服务流。

大模型的发展为智能体生态的发展与协作提供了可能，依托鸿蒙系统，可实现系统与智能体之间、应用与智能体之间、系统智能体与智能体之间的深度融合与协同，带来全新的交互范式与用户智慧化体验升级。



图 5：鸿蒙智能体与系统交互深度融合

鸿蒙系统中构建应用智能体的核心优势

- 丰富的系统触达入口：智能体可以通过丰富的系统级入口触达用户，包括小艺系统智能体、系统导航条、小艺 APP、小艺搜索、小艺建议等。
- 统一的应用智能体协同交互体验：全局体验一致，降低用户认知成本。
- 跨应用智能体协同，闭环服务链：系统智能体负责协同各应用智能体帮助用户最优路径完成任务。
- 1+8 全场景设备协同：满足用户在不同场景、设备下的智能体体验，未来可支持手机、平板等形态，PC、手表等。

3.1.1 智能体全新分发入口

系统级入口

用户通过鸿蒙系统的系统导航条、小艺搜索、小艺建议等系统入口，快速触达相关的鸿蒙智能体，实现随时可达的智能化服务。

◎ 系统导航条：无处不在的系统级全局 AI 入口

系统智能体可通过环境、应用、用户行为等因素，命中智能体或服务意图，并通过系统导航条主动提醒。当系统检测到特定条件时，系统导航条主动提醒智能体；办公、学习、出行等场景下，系统智能体判断对应场景，优先推荐场景相关的服务。

◎ 小艺搜索：桌面下拉快速搜索应用智能体，为用户推荐智能体热度榜单，最热智能体快速一览，方便直达。智能体搜索直达，即搜即得。

◎ 小艺建议：基于情景感知、用户画像能力，联合鸿蒙智能体，基于用户行为，提供智能化主动服务。如订阅博主总结简报、关注话题更新、自动生成旅游攻略等。

小艺系统智能体入口

小艺是鸿蒙系统中的超级智能体，通过便捷的系统级交互（复制、拖拽、圈选等），小艺能基于屏幕内容和用户意图，深度理解用户需求，通过规划拆解任务，协同多个智能体最优路径完成用户任务，提供更加高效、个性化的智慧体验。

应用内入口

通过统一的服务标识和能感知上下文的轻量入口，将智能体从需要用户主动发现和调用的独立功能，转变为与应用核心体验无缝融合、按需涌现的服务。这有助于用户建立一致的认知，把 AI 从“需要主动调用的工具”升级为“自然流动的服务”。



图 6：鸿蒙智能体全新分发入口

3.1.2 智能体全新交互范式

自个人计算机诞生以来，GUI 一直是用户和计算机进行人机交互的主要方式。不管是桌面计算机时代的鼠标、键盘、显示器，还是移动互联网时代人们通过触摸屏进行交互，都是典型的 GUI 交互方式。

在智能体时代，基于意图的个性化服务直达对人机交互提出了全新的要求，鸿蒙智能体框架针对智能体交互提供了不同范式。

智能体操控

通过最自然的交流方式，实现高效、便捷的应用功能访问与控制，操控多步骤任务，极大地提升用户体验和生产效率。用户通过文本、语音、图片等 LUI 为主的方式表达需求，智能体理解用户意图并自动导航执行操作，通过针对目标任务的完成情况，给用户反馈结果。

如用户可以通过“帮我下单一箱我上周购买的饮料和两份上个月我买的零食，要求周五下班前送到我家”、“帮我挑选出我过去收藏歌单里爵士风格的音乐重新组成歌单播放”等意图，完成传统应用交互需要大量操作的步骤。

智能体伴随

这种模式的核心在于用户通过表达一个具有后台持续性任务的意图，智能体基于此意图，从“单次响应”模式切换为“长期陪伴、主动服务”的后台伴随状态。如用户在会议时下达“开启会议助手”指令，智能体持续提供支撑，会议中记录、摘要、翻译，会议后自动跟进待办事项。

智能体 GUI 任务

在一些需要依托于 GUI 显示的长时任务中，通过 GUI+LUI 形成模式互补带来更丰富的交互体验。GUI 界面是直观的视觉呈现，让用户能便捷地浏览和操作各类元素；LUI 则能突破视觉局限，通过自然语言提升操控效率。

除预置 GUI 和智能体开发者自定义 GUI 外，鸿蒙智能体框架还支持动态生成式 UI，构建按需出现、灵活适配的 GUI 卡片式体验，以更好的帮助用户与智能体进行协作，完成用户交代的任务。智能体生成式 UI 交互具备以下优点：

- ◎ **动态**：基于用户意图、设备类型与多模态输入数据，根据用户不同的行为路径和应用场景，实时灵活地判断使用最合理、最自然的交互方式（GUI 还是 LUI，还是混合形式），如何丝滑切换，自然融合。同时对于 GUI，构建功能满足、设计规范、样式美观的轻交互 GUI 卡片。
- ◎ **自主**：基于用户数据、偏好习惯，自主选择对用户最友好的交互形态，生成最有个人喜好与特质的 GUI 卡片；基于智能体类别、属性和生态需求，自主生成具有伙伴和生态特征的 GUI 卡片。
- ◎ **自适应**：基于不同意图，生成的 GUI 卡片具有功能、样式和效果的自适应满足；基于不同设备，生成的 GUI 卡片既保持同一业务场景与用户链路的 GUI 体验一致性，又能感受、体现不同设备的差异性；对不同伙伴，生成的 GUI 卡片具有伙伴和生态特征，自适应伙伴在设计和品牌上的诉求；对不同用户，基于用户喜好和习惯，生成具有个人特征、风格倾向的 GUI 卡片，达成千人千面。

多智能体协同

系统智能体与鸿蒙智能体交互，整体遵循“数据最小化”与“任务最优化”两个基本原则。

1) 数据最小化

系统智能体与应用智能体之间的交互数据，系统智能体会根据用户原始请求进行一定的改写，必要的增删必要信息，在满足用户隐私的基础上召唤应用智能体完成任务闭环。

2) 任务最优化

小艺系统智能体对用户的使用场景和需求理解最深，通过召唤满足独立闭环任务的鸿蒙智能体，采用最优的协同方式，协同多个智能体帮助用户完成任务。

如果单一任务闭环存在多个智能体满足要求，系统智能体可提供给用户选择或优先选择用户体验最优鸿蒙智能体为用户提供服务。

示例：

用户期望小艺帮助完成复杂的任务时，如“帮我安排一场周末的朋友聚会”，系统智能体优先与用户完成沟通和细节需求澄清，然后任务规划，识别需多个智能体协同完成用户任务，最后通过协同多个智能体完成用户请求，提供更加智慧的体验。

3.2 智能体协议升级

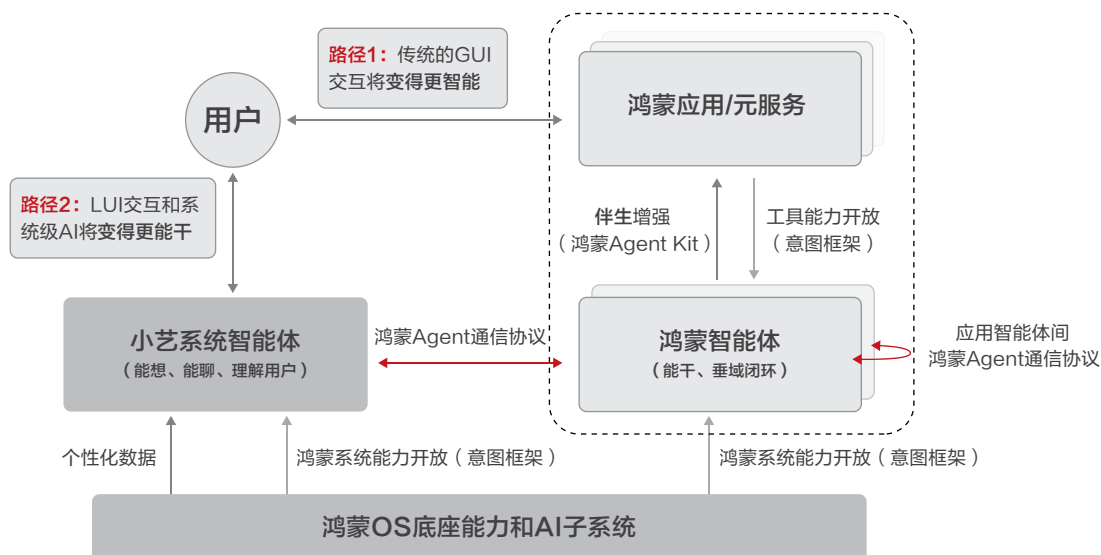


图 7：鸿蒙智能体框架相关协议使用场景

随着应用生态的蓬勃发展，鸿蒙系统智能体与生态应用智能体将构建起完整的智能体互联协作网络，为了更好解决鸿蒙系统下智能体之间高效、可信的协作通信，我们提出了鸿蒙 Agent 通信协议。同时我们对鸿蒙意图框架规范进行升级，更好地解决鸿蒙操作系统和应用 / 元服务为智能体提供工具能力的问题。

3.2.1 鸿蒙 Agent 通信协议实现多智能体高效协同

鸿蒙 Agent 通信协议定义了鸿蒙系统下，小艺系统智能体与应用智能体，以及应用智能体之间的交互规范。主要遵循**五大关键设计原则**：

- 充分对等授权**：鸿蒙 Agent 通信协议构建系统智能体与应用智能体之间的对等交互，基于用户的意图和场景下的充分共享上下文信息，使能应用智能体能够更加精准理解用户意图、闭环复杂任务。
- 生态协议兼容**：鸿蒙 Agent 通信协议充分兼容业界开源的 Agent2Agent 协议，支持三方智能体按照标准 Agent2Agent 协议互通，底层通信协议复用现有的 HTTP、SSE、Streamable、JSONRPC 等业界存量已有的协议框架。在此基础上，也提供了面向鸿蒙系统的扩展增强字段。

3. **安全可靠**：鸿蒙 Agent 通信协议支持企业级安全认证和鉴权机制，智能体之间采用 OAuth2、AK/SK、API Key 实施接口间的通信认证机制，同时支持基于华为账号一键授权登录的用户级认证机制，全链路采用 TLS1.2 及以上版本协议加密传输。除了链路安全外，鸿蒙 Agent 通信协议更注重在智能体业务侧采取满足高等级隐私保护的机制管理隐私数据，比如智能体与应用交互场景下如涉及敏感操作需要用户确认后才能继续执行，智能体在信息不足时主动向用户发起追问澄清。
4. **支持端云复杂任务**：除了云侧信息通信外，部分复杂任务需要进行端侧工具调用，鸿蒙 Agent 通信协议支持长达数小时、需要端云协同运行的复杂任务，用户可以实时感知这些复杂任务的执行状态、中间结果。
5. **全场景跨设备互通**：基于鸿蒙 Agent 通信协议，生态智能体可以做到一次接入，在华为终端全场景 1+8 设备上上进行多设备分发。同时不同设备上的同类型智能体也能充分协同工作，共享上下文。



图 8：鸿蒙 Agent 通信协议使用流程

鸿蒙 Agent 通信协议和核心工作原理如下：

1. **智能体能力发现和声明**：智能体可以定义“智能体卡”来声明其能力范围，从而使系统智能体和其他生态智能体能够充分识别和发现能执行任务的最匹配智能体，智能体的声明支持动态调整，鸿蒙应用智能体通过小艺开放平台来声明其“智能体卡”。
2. **多智能体任务管理模式**：鸿蒙智能体之间的协作支持主从模式、对等协商模式和任务订阅模式等不同模式，鸿蒙 Agent 通信协议支持智能体之间按任务管理模式协同。

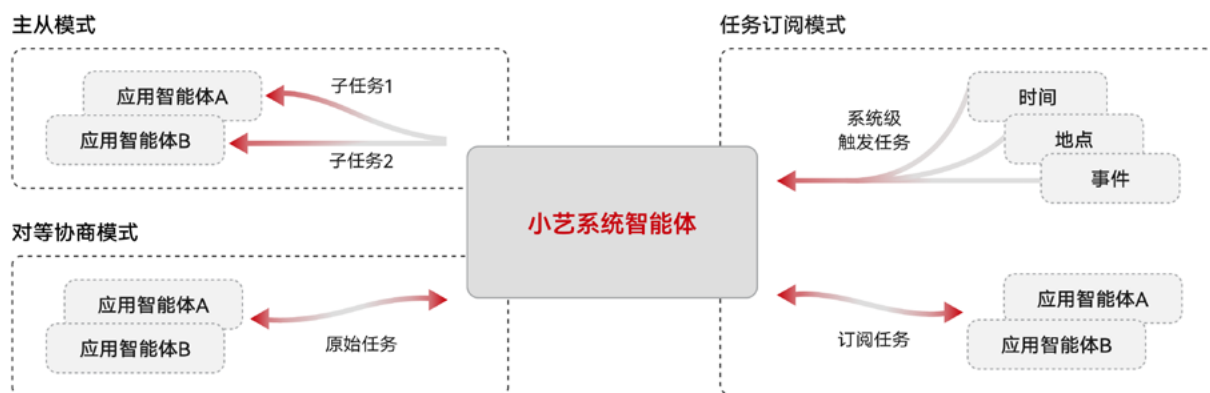


图 9：小艺系统智能体和应用智能体协作的三种模式

主从模式：小艺系统智能体将用户请求自主拆分成不同的子任务，交给不同的协作应用智能体处理，各个应用智能体返回处理结果给小艺系统智能体，小艺系统智能体进行融合答复后返回给用户。在采用鸿蒙智能体通信协议对接的基础上，协作应用智能体既可作为用户可以独立使用的智能体，又可以作为小艺系统智能体处理子任务的智能体。

对等协商模式：分发智能体将用户请求分发给不同的协作智能体，协作智能体独立完成任务，并将处理结果经过分发智能体直接返回给最终用户，如果协作智能体无法完成任务，则将用户请求交还给分发智能体，由分发智能体重新选择其它协作智能体处理。

任务订阅模式：用户通过小艺系统智能体可以给应用智能体下发订阅任务，如国庆节的前一周在应用智能体 A 上帮我预约到北京的机票，每天早上到公司的时候使用应用智能体 B 帮我打上班卡，下午离开公司的时候使用应用智能体 B 帮我打下班卡，小艺系统智能体会基于系统级的时间、地点、事件触发任务管理在满足触发条件时，将用户下发的任务发送给指定的应用智能体处理。

3. 智能体多模态通信：鸿蒙 Agent 通信协议除了允许鸿蒙智能体之间进行文本、文件、图片、音频和视频的通信外，还支持适配鸿蒙系统和 ArkUI 的富文本信息和卡片信息展示，以便于智能体之间的交互结果能够在鸿蒙设备上更加及时地交互和展示。

4. 跨设备、跨应用协作：鸿蒙应用智能体按照鸿蒙 Agent 通信协议开发，与小艺系统智能体对接后，可以支持小艺系统智能体跨设备和跨应用分发，不需要应用智能体做额外改造，跨应用分发依赖小艺系统智能体的跨 APP 调度底层能力，跨设备分发依赖小艺系统智能体的 AI 总线通信能力。



图 10：小艺系统智能体和应用智能体跨设备协同

3.2.2 意图框架全面升级丰富智能体工具生态

智能体生态的成功强依赖于其可使用工具生态的丰富性。好用的工具越多，智能体提供的功能越强大，用户体验越好。为了让智能体的开发者更容易的获取工具，更加便捷的开发好用的工具，解决工具碎片化的问题，鸿蒙意图框架协议在支持标准意图接入的基础上全面升级，新增支持“大模型亲和”的自定义意图规范同时兼容 MCP 协议。在协议升级基础上开放更多的鸿蒙系统工具和 MCP 工具供智能体使用。



图 11：意图框架协议全面升级

升级点 1：支持“大模型亲和”自定义意图的工具开发接入方式，提升接入效率。

开发者可以在代码上添加装饰器完成 APP 内注册，暴露功能给意图框架进行分发。装饰器的内容包括意图（工具）描述、参数描述、示例等均以自然语言的方式开发者自定义，更加符合大模型时代工具的定义方式，相应描述定义的越清晰大模型理解的越好，用户使用体验越好。

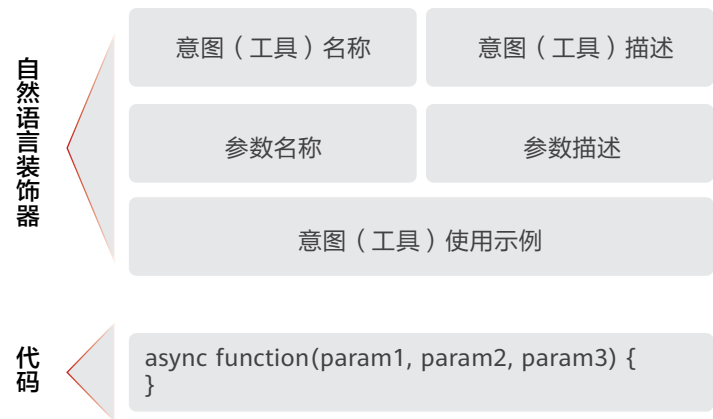


图 12：意图框架自定义意图开发

升级点 2：兼容 MCP 协议，支持 MCP Server 上架、运行和分发。

意图框架协议和 MCP 协议的本质都是作为智能体和工具的连接桥梁，MCP 协议作为行业通用协议，当前已经有较多好用的 MCP 工具，意图框架兼容 MCP 协议，可以让 MCP 工具获得更多的分发入口，同时让智能体有更多的工具选择。按照 MCP 协议开发好的工具，可以免开发接入意图框架并在小艺和智能体中进行分发。

未来意图框架会逐步完善插件工具的管理机制，比如安全管控机制、分发机制等，保证各种插件工具共同协作，有序触达用户。

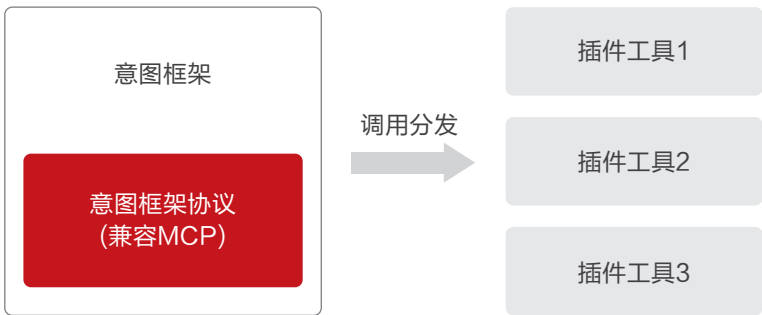


图 13：意图框架兼容 MCP 协议

升级点 3：开放更多工具到智能体插件市场，丰富智能体工具供给。

结合鸿蒙系统和设备的优势，提供更多优质的系统工具，包括但不限于设备操控、系统操控、社交通讯等基础能力，可以让智能体与系统深度融合。同时，支持 MCP 工具发布到智能体插件市场，供智能体开发者使用，增强智能体交互体验。

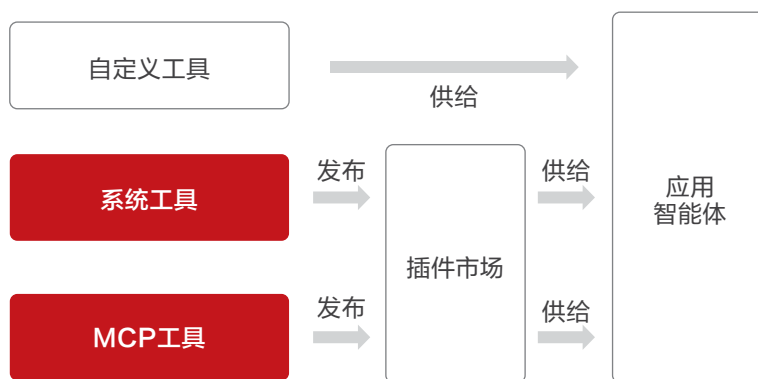


图 14：意图框架通过插件市场为智能体提供工具

3.2.3 鸿蒙 Agent Kit 使能智能体和应用伴生增强

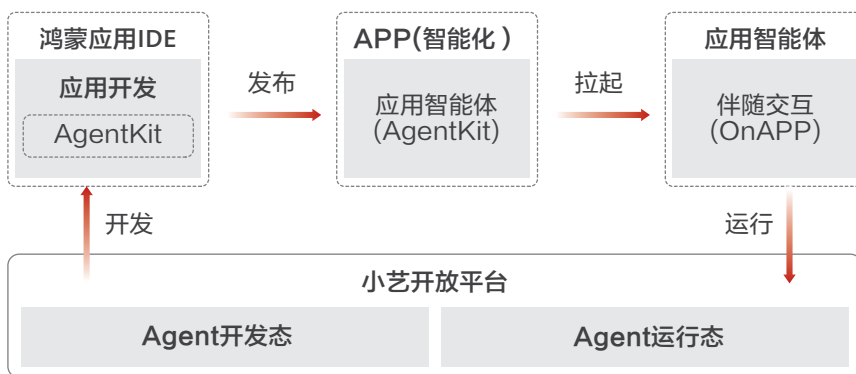


图 15：鸿蒙 Agent Kit 使用流程

鸿蒙 Agent Kit 为鸿蒙应用及元服务提供集成对接鸿蒙智能体的核心能力。开发者通过小艺开放平台完成智能体的开发、编排与上线后，可借助鸿蒙应用开发 IDE 工具 DevEco Studio 在鸿蒙应用和元服务的开发过程中调用鸿蒙 Agent Kit，从而实现应用内智能体的快速集成。

鸿蒙 Agent Kit 的设计原则如下：

1. **组件标准化**：通过统一智能化组件库封装底层能力，规范接口与参数标准，避免跨应用重复开发；
2. **体验一致性**：定义应用内智能化统一风格（UI 反馈、操作逻辑），确保用户跨应用使用智能体时认知一致。
3. **运行标准化**：统一底层运行规则，规范智能体初始化、执行、交互、反馈及资源释放全流程，保障跨应用 / 设备任务逻辑一致。

当鸿蒙应用开发完成并部署安装后，用户在使用鸿蒙应用或者相关应用时，鸿蒙 Agent Kit 支持系统入口、当前应用和相关应用通过主动提醒、被动拉起等交互方式使用鸿蒙应用智能体，实现应用 / 元服务与应用智能体的协同交互。

3.3 智能体高效开发



图 16：小艺开放平台支撑智能体全流程开发、上架管理

小艺开放平台依托鸿蒙系统生态，为开发者打造全链路智能体开发解决方案：平台提供 LLM 模式（大语言模型驱动）、工作流模式（可视化流程编排）、A2A 模式（直连三方智能体）、多 Agent 模式（多智能体协同决策）四大核心开发模式。平台配备端到端工具链，覆盖从智能体开发、多端调试（手机 / 平板 / 车机 / PC / 手表）到部署上架的全生命周期。

智能体完成开发后，将统一上架至智能体市场，实现集中管理和多设备、多入口分发；

为深度赋能开发者，平台推出两大辅助开发能力：

- ◎ AI 辅助开发：自然语言交互生成智能体，降低智能体开发门槛；
- ◎ 鸿蒙 IDE（DevEco Studio）协同：开发者可在同一 IDE 内完成插件、智能体开发、多设备调试，实时调试智能体在手机、车机等不同设备上的运行效果，大幅提升跨端开发效率；

3.3.1 四种开发模式支撑智能体灵活构建

1) LLM 模式

LLM 模式是一种基于大模型的智能体编排方式。开发者按需选择大模型，根据业务逻辑编写提示词，以 LLM 为理解中枢，结合意图识别、工具调用、对话上下文，动态选择插件、工作流，响应用户需求。LLM 模式适用于简单对话、知识问答、基础内容生成等场景。

2) 工作流模式

工作流模式是一种基于规则化流程的智能体编排方式。开发者将复杂任务拆解为有序的规则化步骤（如数据获取、处理、执行），串联插件、大模型、条件分支、代码块等组件实现自动化执行流程，完成业务逻辑。工作流模式适用于需多步骤协同、逻辑复杂、业务多样性的场景。

3) A2A 模式

A2A 模式是一种三方智能体接入小艺开放平台的高效编排方式。开发者可通过该模式基于鸿蒙 Agent 通信协议快速、便捷地将成熟的第三方智能体对接至小艺开放平台，实现分发与调用，提升平台的场景覆盖能力。该模式适用于同时具备鸿蒙端应用与云侧智能体能力的企业开发者。

4) 多 Agent 模式

多 Agent 模式是一种通过编排多个独立智能体协作完成复杂任务的编排方式。开发者可根据业务需求，编排多个鸿蒙应用智能体的功能协作，实现跨应用 / 元服务的工具链调用，完成复杂任务的闭环。多 Agent 模式典型场景包括智能客服分流，企业虚拟员工协作等。

3.3.2 插件市场提供鸿蒙系统级插件

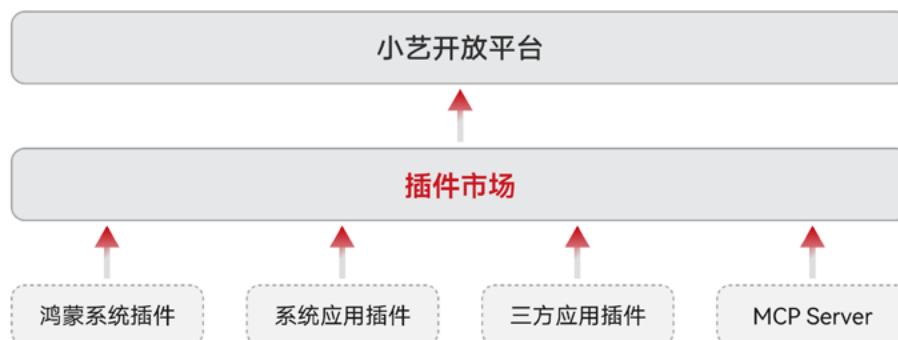


图 17：插件市场支持多种类型工具插件

小艺开放平台通过插件市场实现生态能力的统一接入与开放：平台将鸿蒙系统插件、系统应用插件、第三方应用插件及三方生态 MCP 工具等能力，经可靠性、稳定性测试后，通过开发平台向开发者开放。具体接入范围覆盖以下四类：

- ◎ **鸿蒙系统插件**：鸿蒙 OS 底层能力接口，包含 AI、通信、UI、分布式等核心系统能力；
- ◎ **系统应用插件**：鸿蒙系统应用（如相机、图库、备忘录、闹铃、设置、计算器等）的功能开放接口；
- ◎ **第三方应用插件**：鸿蒙应用市场上架应用的通用能力接口（如社交、工具类应用的功能开放）；
- ◎ **三方生态 MCP 工具**：通过 MCP 协议接入的第三方生态工具能力（三方 saas 服务等）。

通过意图框架开放端云插件能力，支持端+云插件灵活编排，实现：高性能（本地实时交互）、安全可信（本地存储+加密+权限管控）、多端无缝联动（分布式协同），解决高延迟、隐私泄露、跨端易用性的痛点，塑造智能体极致交互体验。

3.3.3 智能体市场高效上架分发



图 18：鸿蒙智能体市场

小艺开放平台通过智能体市场推动鸿蒙智能体生态繁荣，借助精准分发、质量保障、监控运营等手段构建了开发者与终端用户的良性循环：鸿蒙生态开发者完成智能体开发与测试后，可一键发布至智能体市场，获得高效、精准的分发渠道，提升作品曝光与用户触达；终端用户依托严格的质量保障与持续更新机制，始终享有高标准使用体验。

3.3.4 IDE 协同开发

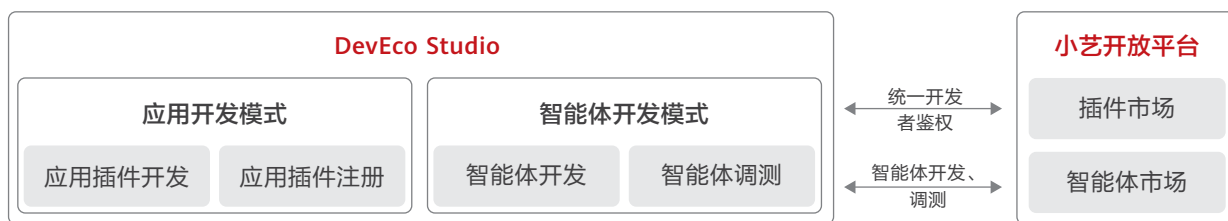


图 19：鸿蒙应用开发 IDE 与小艺开放平台协同开发

小艺开放平台协同鸿蒙 IDE 提供一站式智能体开发环境，支持应用插件与智能体联动开发模式，覆盖从应用插件开发到智能体调测的全流程本地化操作：

- 1) **应用开发模式**：IDE 内置自动生成意图注解能力，开发者通过框选代码快捷操作生成意图注解，快速完成应用内功能向意图框架的开放；生成的意图注解支持一键注册至 Agent 平台，大幅简化端意图暴露流程。

- 2) **智能体开发模式**：IDE 深度集成 Agent 开发框架，统一开发者账号认证鉴权。开发者可在 IDE 内完成智能体全生命周期开发（含已注册端意图插件引用），并支持本地调测，实现从开发到验证的高效闭环。

3.3.5 AI 辅助智能体开发

AI 辅助开发主要是通过与开发者间对话式交互，让开发者在自然语言交互过程中逐步完善需求，动态调整方案，生成满足用户实际需要的 Agent，从而降低构建 Agent 门槛。主要提供 LLM 模式、Workflow 模式两种生成模式。

- ◎ **LLM 模式辅助开发**：根据开发者描述的场景，系统会主动与用户对话澄清需求并动态调整方案。最终构建一个集成插件、知识库、工作流等要素满足用户需求的 Agent。
- ◎ **工作流模式辅助开发**：针对复杂流程任务，用户提供需求描述后，系统将自动构建流程示意图，并配以详细的步骤说明，帮助用户据此交互进行迭代修改。最终基于流程示意图生成直接可执行的 Workflow。



3.4

智能体安全可信

随着智能体技术的快速发展，与现实世界存在诸多恶意诈骗、恶意诱导类似，智能体的功能越强大，它潜在的风险也越大，攻防技术需要不断进化与演进。基于可信基础设施的硬件可信根，华为鸿蒙智能体通过数据安全、模型安全、部署安全、应用安全、运营安全等防线，构建鸿蒙智能体全链路安全体系。

3.4.1 典型 AI Agent 攻击路径

- ◎ **提示词注入攻击**：通过构造恶意智能体提示词，绕过安全限制或诱导 AI 执行非预期操作；
- ◎ **对抗样本攻击**：在输入数据（文本、图像、语音）中添加细微扰动，误导 AI 模型的判断；
- ◎ **权限与工具调用攻击**：利用 AI Agent 被赋予的过度权限执行危险操作，导致鸿蒙系统 Agent 被控制、数据被篡改或窃取；
- ◎ **通信与协议攻击**：劫持 AI Agent 与外部服务（如 API、数据库）之间的通信，利用 AI Agent 交互协议（如 MCP、A2A）设计缺陷，对任务流程被劫持、系统被入侵；
- ◎ **逻辑与任务攻击**：通过误导性指令或分步提示，诱导 AI 执行多阶段恶意任务，通过大量请求或复杂任务耗尽 AI Agent 的计算资源；
- ◎ **供应链与依赖攻击**：在模型训练阶段注入恶意样本或参数，利用 AI Agent 依赖的第三方服务（如云平台、API）的漏洞等。

3.4.2 被保护关键资产

在鸿蒙智能体关键资产梳理中，有以下最关键的四类关键资产：

用户敏感数据：如个人身份信息（PII）、支付信息、生物特征等，通过加密或脱敏保护。

设备访问权限：包括摄像头、麦克风、位置等硬件权限，防止未经授权的调用。

用户身份凭证：如账号密码、OAuth 令牌、用户标识等，应由安全模块存储并严格控制访问；

基础设施系统与应用配置信息：如 API 密钥、网络设置等，防止被恶意读取或修改；

3.4.3 对应的典型防御阵线

结合鸿蒙智能体需要保护的关键资产、典型的攻击路径，鸿蒙智能体框架 HMAF 在传统软件安全工程的基础上，迭加构建 AI Agent 典型安全防线，构建基于华为可信基础设施的全栈安全防线：

1. **可信基础设施：**华为提供基于硬件的可信根安全能力，包含设备身份、设备密钥派生、密钥管理、设备证书链等，提供硬件级密钥的安全能力、安全运算环境，保障上层业务敏感信息的机密性、完整性。
2. **智能体数据安全：**提供基于设备硬件的智能体数据加密机制，依托 TEE(可信执行环境) 或专用安全芯片实现全生命周期加密，同时增加锁屏下数据分级访问机制，确保系统智能体数据在锁屏后无法被恶意程序访问。
3. **智能体端云执行环境安全：**基于设备级密钥协商机制、盲化链路、云上硬件级可信根，构建跟用户设备独有的私有计算区，只有用户自己的设备能够解密，且任意个人 / 厂商，包括华为，无法触碰私有计算区的数据。
4. **智能体统一身份认证：**基于华为账号统一身份的 OAuth 2.0 认证机制，让智能体资源调用、智能体间调度时保障跨 Agent 的数据资产不被恶意访问、篡改，确保只有自己的账号能访问你的 Agent。
5. **智能体运营安全：**基于统一账号的 Agent 安全运营，对恶意 Prompt、恶意攻击、恶意 Agent 持续监控，通过信用分机制对 Agent 主体、账号个人进行评分，不断提升 Agent 合法合规运行能力。

A scenic landscape featuring a hot air balloon with a white, yellow, and orange striped pattern floating in the sky. Below the balloon is a calm lake nestled in a valley, surrounded by rugged mountains. The foreground shows a rocky, grassy hillside. The sky is filled with soft, colorful clouds, suggesting a sunrise or sunset. The overall mood is peaceful and expansive.

04

总结与未来展望

截止 2025 年 6 月，已有超过 800 万的开发者投入到鸿蒙生态的开发中来，基于鸿蒙智能体框架（Harmony Agent Framework, HMAF），通过小艺智能体开放平台和鸿蒙系统能力将充分赋能开发者，使能智能体的快速构建和部署，催生“智能体即服务”（AaaS）的商业模式。鸿蒙应用生态和智能体技术将深度结合，通过面向智能体的系统级全新交互、高效可信的通信协议、开发能力和安全体系，鸿蒙智能体将全面赋能鸿蒙应用智能化演进，快速提升各个垂域的智能服务体验，形成“应用 + 智能体”的共生生态。

基于鸿蒙智能体框架，全场景终端设备、鸿蒙操作系统、鸿蒙应用和智能体将构成一个充分协同的 AI 生态体系，形成高效、自主和可控的人机交互新范式，支持“以指令为中心”的交互向“以意图为中心”升级，打通用户、系统与服务之间的链路，带来全新的服务模式，有效提升华为终端消费者在办公、家庭、出行等不同场景中体验。过去通过人进行 APP 操作串联，人记不住复杂的操作步骤，未来可以通过自然语言交代，智能体与应用协同完成。

最后，我们呼吁更多的开发者、企业和个人加入鸿蒙 AI 生态，拥抱人工智能和智能体带来的机遇和挑战，携手共筑一个开放、创新、合作的智能世界！

05

参考资料

- [1] 《AI 终端白皮书》，<https://developer.huawei.com/consumer/cn/doc/guidebook/terminals-ai-0000001929853698>
- [2] 《鸿蒙 2030 白皮书》：<https://developer.huawei.com/consumer/cn/doc/guidebook/bps2023-0000001957043797>
- [3] “Personal LLM Agents: Insights and Survey about the Capability, Efficiency and Security” , Institute for AI Industry Research (AIR), Tsinghua University, <https://arxiv.org/abs/2401.05459>
- [4] <https://www.gartner.com/en/articles/top-technology-trends-2025>
- [5] <https://www.grandviewresearch.com/industry-analysis/ai-agents-market-report>
- [6] <https://www.deloitte.com/cn/zh/Industries/tmt/perspectives/ai-model-c-end-ecology.html>
- [7] <https://github.com/modelcontextprotocol>
- [8] <https://github.com/google-a2a/A2A>

华为终端有限公司

地址：广东省东莞市松山湖园区新城路2号

网址：<https://consumer.huawei.com>

版权所有 ©华为终端有限公司 2025。保留一切权利。

本材料所载内容受著作权法的保护，著作权由华为公司或其许可人拥有，但注明引用其他方的内容除外。未经华为公司或其许可人事先书面许可，任何人不得将本材料中的任何内容以任何方式进行复制、经销、翻印、播放、以超级链路连接或传送、存储于信息检索系统或者其他任何商业目的的使用。

商标声明



以上为华为公司的商标（非详尽清单），未经华为公司书面事先明示许可，任何第三方不得以任何形式使用。

注意

华为会不定期对本文档的内容进行更新。

本文档仅作为使用指导，文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。