

火山引擎AI相关五大核心平台详细解析

瞎聊产品 2026年2月6日 20:07 上海



随着AI技术进入规模化落地阶段，企业级AI Agent成为推动数字化转型的核心动力，云与AI基础设施、大模型服务、智能体开发工具的协同能力，成为科技企业核心竞争力。火山引擎作为字节跳动旗下云和AI服务平台，整合内部技术沉淀，推出火山引擎（基础平台）、火山方舟、AgentKit、HiAgent、Coze五大核心平台，构建从基础设施到应用落地的全链路AI生态。本次调研聚焦五大平台，梳理其官方定位、核心功能、目标用户及相互协同关系，为企业选型、行业研究提供客观参考。



1-火山引擎（基础云与AI平台）



AI云原生基础设施底座
聚焦GPU算力集群与模型定制

核心功能矩阵

- 高性能算力：**提供AI云原生基础设施，支持模型深度调优与定制。
- 安全合规：**自研防火墙体系，全方位保障企业数据安全与内容合规。

核心竞争优势

- 极致性价比：**依托字节跳动规模优势，实现同款GPU配置价格最优。
- 深厚技术沉淀：**经过内部50+业务场景大规模验证，提供全栈一体化支撑。

公众号 · 瞎聊产品

官方定位

字节跳动旗下云和AI服务平台，将字节跳动快速发展过程中积累的增长方法、技术能力和应用工具开放给外部企业，通过云和智能技术帮助企业构建体验创新、数据驱动和敏捷迭代等能力，推进企业AI转型，激发增长潜能，定位为整个生态的底层基础设施底座。2024冬季FORCE原动力大会上，火山引擎正式提出计算范式将从云原生进入到AI云原生的新时代，聚焦提供以GPU为核心的AI云原生基础设施。

核心功能

AI云原生基础设施：提供以GPU为核心的算力集群、高速互联带宽，支持全栈系统化推理优化，具备模型调优和一站式定制能力；

全栈云服务：涵盖弹性计算、对象存储、容器服务、网络服务等基础云能力，支撑上层平台的稳定运行；

安全与合规体系：自研防火墙，可提升算力可用性、降低数据泄露风险、保障内容合规；

定制化服务与技术支持：提供1V1架构师定制方案、7x24小时人工客服、飞书值班号快捷响应，以及专业技术工程师支持，实现落地闭环。

目标用户

各类有云服务、AI基础设施需求的企业客户，重点是IT团队、技术负责人，覆盖互联网、汽车、金融、制造、教育等30+行业，包括初创企业、中型企业及大型集团。

核心优势

性价比突出：得益于与字节跳动国内业务并池的规模优势，可实现同款GPU配置价格最优；

技术沉淀深厚：依托字节内部50+业务场景实践验证，算力稳定性、推理效率经过大规模场景打磨；

全栈支撑能力：从基础算力到安全合规，提供一体化基础设施服务，无需企业额外整合第三方资源；

服务体系完善：定制化方案+全方位响应+专业化技术支持，降低企业落地门槛。

产品动态

2024冬季FORCE原动力大会上，明确提出AI云原生战略，升级GPU为核心的基础设施；持续优化算力调度能力，提升与上层火山方舟、AgentKit等平台的协同效率；不断拓展行业覆盖，目前已服务30+行业的众多客户。

2-火山方舟（大模型服务平台/MaaS）



AI控制中心与模型服务枢纽

📍 官方定位：一站式 MaaS 平台

定位为“AI控制中心”，是连接大模型与企业应用的核心枢纽，为上层智能体平台提供核心能力。

🔧 核心功能：全链路模型服务

汇集豆包家族及第三方主流模型，支持多维度选型对比；提供在线推理、精调及API调用服务，整合插件生态保障数据安全。

🏆 核心优势：生态丰富与易用性

整合自研与第三方模型，覆盖多类型需求；可视化对比简化选型流程，与底层算力及上层智能体无缝对接，降低企业使用门槛。

🗣️ 公众号 · 瞎聊产品

官方定位

火山引擎一站式大模型服务平台（MaaS，Model as a Service），定位为“AI控制中心”，汇集时下主流大模型，提供模型训练、推理、评测、精调全流程服务，实现大模型的统一管理与高效复用，为上层智能体开发平台提供核心模型能力支撑，是连接大模型与企业应用的核心枢纽。

核心功能

模型广场：汇集多类型、多厂商大模型，包括豆包家族全系列模型（语言、视觉、语音、向量模型），以及DeepSeek、百川智能、出门问问等第三方模型，支持列表视图与厂商视图切换；

模型管理与对比：支持以模型效果、能力、价格、tokens限制、任务类型等维度对比模型，提供详细模型卡片与说明文档，方便用户快速选型；

模型服务能力：提供在线推理、模型精调、API调用等服务，支持自定义推理接入点，可快速为AgentKit、HiAgent、Coze等平台提供模型能力输出；

插件生态与安全保障：整合各类插件资源，结合火山引擎底层安全能力，保障模型调用与数据传输的安全合规。

目标用户

算法工程师、技术开发者、企业IT团队、产品团队，核心是需要快速使用大模型能力、无需自行搭建模型训练与推理环境的用户，覆盖需要模型精调、推理部署的各类企业场景。

核心优势

模型生态丰富：整合豆包自研模型与第三方主流模型，覆盖语言、视觉、语音等多类型，用户可按需选型，无需绑定单一模型；

易用性强：提供可视化模型对比、详细说明文档，简化模型选型与调用流程，降低大模型使用门槛；

协同能力突出：与火山引擎底层算力深度协同，保障模型推理效率；与AgentKit、HiAgent、Coze无缝对接，实现模型能力快速输出；

技术支撑有力：依托豆包大模型的领先性能（智源等权威机构测评国内领先），为用户提供高质量模型服务。

产品动态

持续扩充模型广场的模型数量与类型，新增多厂商第三方模型；优化模型对比与选型功能，提升用户体验；强化与上层智能体开发平台的协同，支持模型精调后快速部署至AgentKit、HiAgent、Coze等平台；同步豆包大模型的版本更新，持续提升模型服务能力。

3-AgentKit（企业级AI Agent基础设施底座）



官方定位

火山引擎企业级AI Agent基础设施底座，定位为降低复杂AI Agent开发门槛的核心技术支撑平台，提供Agent全生命周期开发所需的底层技术模块，支持开发者高效构建稳定、安全、可扩展的企业级AI Agent，是上层HiAgent平台的核心技术支撑之一，也是火山引擎智能体生态的技术基石。

核心功能

开发环境与工具：支持Python和Golang开发，提供多种模板（基础应用、流式输出、A2A协议等），简化Agent开发流程；

模型对接能力：深度对接火山方舟大模型服务平台，支持调用火山方舟中的各类模型（含豆包及第三方模型），可自动注入内置豆包模型，也支持手动配置模型推理API Key与Endpoint；

底层技术支撑：提供Agent开发所需的核心技术模块，包括身份认证、运行时管理、云沙箱、网关等，保障Agent的稳定运行与安全合规；

部署与调试：支持本地调试与云端部署，提供访问密钥配置功能，首次部署可快速完成（约2-3分钟），方便开发者快速验证与迭代Agent产品。

目标用户

企业技术团队、高级开发者、算法工程师，核心是需要构建复杂企业级AI Agent、具备一定编程能力，追求Agent定制化与稳定性的技术从业者，覆盖需要深度定制AI Agent的大型企业与技术驱动型企业。

核心优势

开发效率高：提供丰富模板与简洁的开发接口，大幅降低复杂Agent的开发难度，官方数据显示可使Agent开发代码量减少96%；

技术稳定性强：依托火山引擎底层基础设施与安全体系，提供全链路监控与安全保障，确保Agent稳定运行；

协同性好：与火山方舟深度协同，可快速调用各类大模型能力；与HiAgent无缝对接，开发完成的Agent可直接接入HiAgent进行运营管理；

定制化能力突出：提供底层技术模块的灵活配置，支持开发者根据企业业务需求，构建高度定制化的企业级AI Agent。

产品动态

2025年12月FORCE原动力大会上全面升级，强化与火山方舟、HiAgent的协同能力；持续优化开发模板与工具，支持更多开发语言与场景；完善本地调试与云端部署功能，提升开发者体验；新增安全围栏等模块，进一步强化Agent的安全合规能力。

4-HiAgent（一站式智能体工作站）



企业数字员工管理中枢

🔗 “1+N+X” 全栈智能体体系

- 构建统一交互入口与通用智能体中心，支持多Agent协同纳管
- 集成Agent DevOps，覆盖从开发、运维到优化的全生命周期管理

🔄 全流程闭环与低代码开发

- 国内首个实现“开发-运营-使用”全闭环，解决规模化落地难题
- 提供低代码/高代码混合模式，大幅降低业务人员使用门槛

👤 广泛的用户覆盖与行业落地

- 服务于业务人员、IT团队及管理者，满足规模化运营AI Agent需求
- 已在众多头部客户落地，产品能力处于行业领先地位

🗨️ 公众号 · 瞎聊产品

官方定位

火山引擎企业级一站式智能体工作站，国内首个实现智能体从生产、运营到使用的全流程闭环体系化平台，定位为“企业数字员工管理中枢”，基于Agent DevOps理念，提供智能体开发、评测、观测、优化全生命周期管理，构建“1+N+X”体系，实现智能体的规模化运营与协同，助力企业将AI Agent转化为数字生产力。

核心功能

“1+N+X”体系：1个统一交互入口（AgentSphere），实现所有智能体的统一调度与人机协作；N个通用智能体中心，提供开箱即用的办公自动化、写作、审核等通用智能体；X个开发运维平台，实现智能体全生命周期管理；

Agent全生命周期管理（Agent DevOps）：涵盖策略规划、能力开发、效果评测、应用发布、线上观测、品质优化全流程，实现智能体从“实习生”到“职场老员工”的持续进化；

智能体协同与纳管：支持多智能体协同调度，可纳管企业内部存量智能体及第三方开发的智能体，通过A2A+MCP协议低成本集成所有智能体和业务系统；

低代码开发能力：提供低代码与高代码混合开发模式，支持从模板挑选工具，快速搭建“数字员工”，适配大型组织的复杂业务需求；

行业适配与案例落地：提供多行业智能体解决方案，已在金融、教育、零售等行业实现规模化落地，具备完善的效果评测与优化能力。

目标用户

企业业务人员、IT团队、企业管理者，核心是需要规模化运营AI Agent、无需深度编程即可构建业务型智能体的企业用户，覆盖大型集团、中型企业，重点服务金融、教育、零售、制造等行业的头部企业。

核心优势

全流程闭环能力：国内首个实现智能体“开发-运营-使用”全流程闭环，解决企业智能体“散、乱、用不起来”的痛点；

易用性突出：低代码混合开发模式，结合“1+N+X”体系，降低业务人员使用门槛，实现千人千面的人机协同；

行业落地能力强：已服务中国银联、西安交大、周大福、海亮教育等头部客户，在多行业实现规模化落地，IDC报告将其列入中国智能体开发平台领导者象限，产品能力位居中国第一；

协同生态完善：与AgentKit、火山方舟、Coze深度协同，可接入AgentKit开发的定制化Agent、Coze开发的轻量化Agent，整合火山方舟的模型能力，形成完整的智能体生态。

产品动态

2024年首次发布，2025年推出2.0版本，升级“1+N+X”体系，新增AgentSphere统一交互入口与Agent DevOps全生命周期管理能力；2025年12月发布全新版本，强化多智能体协同与纳管能力；

持续拓展行业案例，完善行业模板；与合合信息等合作伙伴联手，丰富插件与工具生态，提升文档解析等场景的能力。

5-Coze（扣子，AI Agent开发工具）



低门槛·场景化·普惠型
全量开发者与中小企业的首选

官方定位：一站式AI Agent开发平台
定位“低门槛、场景化”，与HiAgent形成分层服务体系，专注于快速构建轻量化智能体。

核心功能特性

- 低代码/零代码开发**
可视化编排，无需专业编程经验
- 1万+ 丰富插件生态**
支持多模态数据格式，覆盖多场景
- 多模态与记忆能力**
支持文本/图像处理，具备长期记忆
- 灵活部署与开源支持**
支持公有云及私有化开源部署

核心优势

- 极速搭建**
3天内完成轻量化应用构建
- 成本可控**
开源免费+企业订阅模式
- 生态丰富**
海量插件满足多样化需求

公众号·瞎聊产品

官方定位

火山引擎一站式AI Agent开发工具，定位为“低门槛、场景化”的普惠型智能体开发平台，面向全量开发者与个人用户，支持可视化配置+简单代码编写，快速构建AI应用与智能体，采用“开源免费+企业订阅”模式，与HiAgent形成分层服务体系，覆盖轻量化、快速落地的智能体开发场景，是火山引擎智能体生态的普惠型补充。

核心功能

低代码/零代码开发：提供可视化设计与编排工具，无需专业编程经验即可快速搭建AI智能体，支持模块化、可视化的工作流设计；

丰富插件与知识库：拥有1万+插件，涵盖图像类、文本类、搜索类、数据分析、语音识别等多个领域；提供简单易用的知识库功能，支持多种数据格式与数据源，实现智能体与用户自有数据的交互；

多模态与记忆能力：支持多模态模型，可处理文本、图像等多种类型数据，提供丰富的交互体验；具备长期记忆能力，可持久记住用户对话中的重要参数与内容，提升交互连贯性；

部署与开源支持：支持公有云、开源个人应用等多种部署模式，第三方开发者可通过Coze Marketplace销售自定义插件和智能体模板；提供扣子空间，实现开发者与AI Agent的协同办公；

模型对接：深度对接火山方舟大模型服务平台，默认调用方舟输出的豆包大模型能力，无需自行配置复杂的模型接口，简化开发流程；

目标用户

全量开发者（含个人开发者）、中小企业、业务人员，核心是需要快速构建轻量化AI智能体、无需深度定制，追求开发效率与低成本的用户，覆盖个人创作、中小企业办公、轻量化业务场景等。

核心优势

门槛极低：零代码/低代码开发模式，可视化操作，即使没有编程经验也能快速上手，3天内可搭建多语言客服系统等轻量化应用；

生态丰富：1万+插件覆盖多场景，知识库功能完善，支持多模态交互，可满足多样化的智能体开发需求；

成本可控：开源免费+企业订阅模式，个人开发者可免费使用基础功能，中小企业可按需订阅，降低开发与使用成本；

协同性好：与HiAgent深度协同，开发的轻量化智能体可接入HiAgent进行统一管理与规模化运营；对接豆包大模型，确保核心AI能力的稳定性与高质量。

产品动态

持续扩充插件库数量与类型，完善知识库与 workflows 功能；优化零代码开发体验，提升可视化编排的灵活性；推进开源生态建设，吸引

第三方开发者参与插件与模板开发；与合合信息等合作伙伴联手，丰富文档解析等垂直场景的能力，提升轻量化智能体的落地效率；优化与HiAgent的协同对接，实现智能体的快速迁移与规模化运营。

五大平台相互关系梳理

火山引擎五大核心平台并非孤立存在，而是基于“基础设施-模型服务-开发底座-运营管理-普惠开发”的分层架构，形成逻辑清晰、协同紧密的完整AI生态，各平台各司其职、优势互补，共同支撑企业AI Agent的规模化落地。

1-整体技术分层架构



五大平台分为4个核心层级，自上而下形成完整的能力传递链路，底层平台为上层平台提供核心支撑，上层平台实现底层能力的商业化落地与普惠化延伸，具体分层如下：

底层基础设施层

火山引擎（基础云与AI平台）——整个生态的“地基”，为所有上层平台提供GPU算力、存储、网络、安全合规等基础设施支撑，是火山方舟、AgentKit、HiAgent、Coze稳定运行的核心保障，所

有上层平台的算力、部署均依赖于火山引擎的AI云原生基础设施；

模型服务层（MaaS）

火山方舟——生态的“AI引擎中枢”，整合豆包及第三方大模型，提供模型推理、精调、管理等服务，为AgentKit、HiAgent、Coze三大智能体相关平台提供核心AI模型能力，是连接底层算力与上层智能体开发的关键纽带，AgentKit、HiAgent、Coze的核心AI能力均来自火山方舟；

智能体开发与运营层

AgentKit（开发底座）+ HiAgent（运营管理）——生态的“企业级智能体核心载体”，两者协同覆盖企业级智能体的“开发-运营”全流程，AgentKit提供底层技术开发能力，HiAgent基于AgentKit的技术支撑，实现智能体的全生命周期管理与规模化运营，形成“开发-部署-运营”的企业级闭环；

普惠型开发工具层

Coze（扣子）——生态的“普惠型补充”，面向个人与中小企业，提供低门槛的智能体开发工具，与HiAgent形成分层服务体系（HiAgent侧重企业级规模化运营，Coze侧重轻量化快速开发），其模型能力来自火山方舟，开发完成的智能体可接入HiAgent实现规模化运营，填补普惠型智能体开发场景的空白。

2-各平台具体协同关系



火山引擎与其他四大平台：支撑与被支撑的核心关系

火山引擎是其他四大平台的核心底层基础设施提供商，所有平台均依托火山引擎的AI云原生算力、存储、网络与安全能力，实现稳定运行与高效部署：

为火山方舟提供模型训练、推理所需的GPU算力与全栈系统化推理优化，保障模型服务的高效与稳定；

为AgentKit提供云沙箱、运行时、安全围栏等底层技术支撑，以及云端部署环境，确保Agent开发与运行的稳定性；

为HiAgent提供大规模智能体运营所需的算力、数据存储与业务系统集成能力，支撑多智能体协同与规模化落地；

为Coze提供插件存储、知识库数据存储、云端部署等基础设施，保障低代码开发与快速部署的效率；

同时，火山引擎的安全合规体系贯穿四大平台，保障所有平台的数据安全与内容合规，提供统一的技术支持与服务体系。

火山方舟与AgentKit、HiAgent、Coze：模型能力的传递关系

火山方舟作为大模型服务中枢，是三大智能体相关平台的核心AI能力来源，实现模型能力的统一输出与灵活调用：

与AgentKit: AgentKit通过API对接火山方舟，可灵活调用方舟平台中的豆包及第三方模型，开发者无需自行搭建模型环境，可直接基于方舟的模型能力，开发定制化企业级Agent，AgentKit也可自动注入方舟内置的豆包模型，简化开发流程；

与HiAgent: HiAgent深度集成火山方舟的模型服务，用户可在HiAgent平台直接选用方舟中的模型，用于智能体的开发与优化，模型精调后可直接部署至HiAgent，实现“模型-智能体-运营”的无缝衔接；

与Coze: Coze深度对接火山方舟，默认调用方舟输出的豆包大模型能力，开发者无需配置复杂的模型接口，可直接调用模型能力快速构建多模态智能体，插件生态也与方舟模型能力深度协同，提升智能体的功能丰富度；

核心逻辑: 火山方舟将大模型能力“标准化、服务化”，降低三大平台的模型使用门槛，实现模型能力的统一管理与高效复用，同时同步豆包大模型的版本更新，持续提升三大平台的AI能力。

AgentKit与HiAgent: 开发底座与运营中台的协同闭环

AgentKit与HiAgent是火山引擎企业级智能体生态的核心组合，两者分工明确、深度协同，形成“开发-运营”的完整闭环，解决企业级智能体“能开发、能落地、能运营”的核心需求：

分工协同: AgentKit聚焦“技术开发”，解决“如何高效开发部署稳定可靠的企业级智能体”的技术问题，面向专业技术团队，提供深度定制能力；HiAgent聚焦“运营管理”，解决“如何让智能体在企业中规模化运营、创造业务价值”的应用问题，面向业务人员与企业管理者，提供低代码易用性与规模化能力；

技术协同: HiAgent基于AgentKit的底层技术模块（身份认证、运行时、云沙箱等）构建，保障智能体的稳定运行与安全合规；AgentKit开发完成的企业级智能体，可无缝接入HiAgent平台，通过Hi

Agent的“1+N+X”体系进行统一纳管、协同调度与全生命周期运营；

落地协同：AgentKit提供深度定制能力，满足企业复杂业务场景的智能体开发需求；HiAgent提供运营优化能力，通过Agent DevOps全流程管理，让智能体持续进化，适配业务变化，同时提供行业模板，结合AgentKit的定制化能力，加速企业级智能体的落地效率；

共同目标：两者协同，为大型企业提供“定制化开发+规模化运营”的端到端企业级智能体解决方案，助力企业实现数字员工的规模化上岗，这也是火山引擎入选IDC中国智能体开发平台领导者象限的核心原因之一。

Coze与HiAgent：分层服务与协同互补

Coze与HiAgent均聚焦智能体的开发与运营，但目标用户与场景定位不同，形成“普惠型开发+企业级运营”的分层协同体系，覆盖不同用户群体的需求：

定位互补：Coze面向个人开发者与中小企业，侧重“低门槛、轻量化、快速开发”，解决轻量化智能体的快速落地需求，采用开源免费模式，降低开发成本；HiAgent面向大型企业，侧重“企业级、规模化、全流程运营”，解决复杂智能体的管理与协同需求，采用企业订阅模式，提供专业化服务；

能力协同：Coze开发的轻量化智能体，可通过标准化接口接入HiAgent平台，实现从“个人/中小企业原型”到“企业级规模化运营”的平滑过渡，解决Coze智能体的规模化运营痛点；HiAgent的企业级安全、合规与协同能力，可补充Coze在企业级场景的短板，拓展Coze智能体的应用范围；

生态协同：两者共同接入火山方舟的模型能力与火山引擎的底层基础设施，形成“普惠开发-企业运营”的协同生态，既满足个人与

中小企业的低成本开发需求，也满足大型企业的规模化运营需求，完善火山引擎智能体生态的覆盖范围；

合作案例：合合信息与火山引擎合作，在HiAgent、Coze平台中共同构建文档解析领域的分层服务体系，HiAgent面向企业级用户提供复杂场景解决方案，Coze面向开发者提供低门槛工具，实现全场景覆盖。

五大平台协同核心目标

五大平台协同的核心目标是“降低企业AI Agent的开发、部署、运营门槛，实现AI智能体的规模化落地，助力企业AI转型”，通过分层架构与协同机制，覆盖“基础设施-模型-开发-运营”全链路，满足不同规模、不同技术能力、不同业务场景的企业与开发者需求，构建“技术完善、生态丰富、分层覆盖”的火山引擎AI Agent生态，依托字节跳动的技术沉淀，在企业级智能体赛道形成核心竞争力。

瞎聊！

作者提示: 个人观点，仅供参考

