下载日志已记录，仅供内部参考

计算机

报告日期：2026 年 09 月 22 日

Jev 探索模型新分工，Muse 拓展个人 Agent 生态

——计算机行业点评报告

投资要点

- **Jev 以低成本决策探索模型新分工，拓展高频任务的自动化空间。**

9 月 15 日，TypeSafe AI 发布专用决策模型 Jev，采用面向校准决策的强化学习（RLCD），直接返回结构化判断及概率，为自动处理与人工复核提供依据。通过一次调用并行完成多个独立判断，Jev 减少文本生成与结果解析开销，为高频分类、评分和路由提供低成本基础。其输入价格为 0.042 美元/百万 Token，输出免费，官方工作流测试显示显著成本优势。Vercel 披露，Jev 上线 AI Gateway 后 24 小时内被近 13%的付费团队采用，成为该平台历史上采用速度最快的模型。我们认为，专用决策模型有望与承担复杂推理和内容生成的通用模型形成分工，降低完整任务的处理成本，使更多细小、高频环节具备自动化价值。

- **Meta 发布 Muse，推动个人助理从提供建议走向持续办理事务。**

9 月 8 日，Meta 发布个人 Agent Muse，以自研 Muse Spark 模型为基础，结合云端运行环境、长期记忆与工具调用，在用户关闭应用后仍可推进任务。其能力已覆盖信息整理、文档交付、购物和预订等场景。截至 9 月 20 日，Muse 位居美国两大应用商店免费总榜第一；研究平台 FUNDA 对收集的 654 个公开 Muse 使用案例进行分类分析发现，节省开支、工作与小企业经营、行政事务是占比最高的三类应用场景。我们认为，跨应用执行与持续任务管理，使 Agent 能够直接创造省时、省钱的价值，为用户形成长期使用习惯及付费意愿提供基础。

- **互联网生态与智能终端协同，有望拓展个人 Agent 的服务入口。**

Muse 依托 WhatsApp 等自有入口，并通过 Shopify、Stripe、Plaid 等合作补齐商品、支付和账户数据能力，推动用户需求向实际服务转化。Meta 已提出将 Muse 带入 AI 眼镜，随身感知与语音交互有望降低任务发起门槛。自有社交入口有助于降低获客成本、增加用户触达与任务发起频次，互联网服务生态提供可执行的服务，眼镜等终端进一步拓展使用场景。自有入口、服务生态与端云协同相结合，有望推动个人 Agent 成为日常生活中的服务入口。

- **账户授权与实际执行推动安全需求升级，AI 安全产品迎来拓展空间。**

Muse 通过凭据隔离、独立审批和交易确认，将 Agent 操作限制在用户授权范围内，为连接个人账户和持续办理事务提供基础。随着 Agent 接触更多敏感数据并参与交易，安全需求将从内容过滤延伸至身份认证、权限管理、运行防护与审计追溯。我们认为，可信执行是深化用户委托的必要条件，具备相关产品和客户基础的安全厂商有望受益于 Agent 应用扩张。

- **投资建议**

首先，Jev 类决策模型带来了新的训练范式，将为 AI 办公、coding、AI 安全等智能体提供更快决策速度以及更低成本，**可关注 AI Agent 领域公司金山办公、道通科技、中控技术、金蝶国际等。**

其次，随着个人 Agent 的普及，提供可信环境的网络安全公司需求有望提升，**可关注网宿科技、深信服、绿盟科技、三未信安等安全公司。**

- **风险提示**

应用效果及商业化不及预期；服务接入受限及隐私安全风险；价格竞争与持续运行成本导致盈利能力承压。

行业评级：看好(维持)

分析师：刘雯蜀
执业证书号：S1230523020002
liuwenshu03@stocke.com.cn

分析师：郑毅
执业证书号：S1230524070002
zhengyi@stocke.com.cn

相关报告

1 《开放模型与 Harness 重塑成本结构》 2026.09.19

2 《AI 办公：生产力的代际分水岭》 2026.09.19

3 《DeepSeek 再升级，OpenAI 实现自动化研究实习生》 2026.09.12

下载日志已记录，仅供内部参考

正文目录

1 Jev：低成本决策模型探索 Agent 新分工 4
2 Muse：Meta 自研模型驱动，个人助理开始持续办事 6
3 安全：账户授权与执行防护构成个人 Agent 的基础能力 8
4 生态：互联网服务持续扩展，AI 眼镜打开新的交互入口 8
5 模型分工、服务生态与 AI 安全共同推进落地 9
6 风险提示 10



下载日志已记录，仅供内部参考

图表目录

图 1： Jev 以三类结构化输出承接高频决策 4
图 2： Jev 在官方工作流测试中以较低成本实现接近通用模型的准确率 5
图 3： Jev 上线首日被近 13%的付费 AI Gateway 团队采用 5
图 4： Muse 已登顶美国两大应用商店免费总榜 6
图 5： Muse 上线初期累计下载快速增长 6
图 6： Muse 可将跨应用任务交付为文档、工具和仪表盘 6
图 7： 目标管理、活动记录和确认卡片支撑长期委托 7
图 8： Muse 凭据隔离与独立审批将 Agent 操作限制在授权范围内 8
图 9： 连接器审核与目录分发有助于扩大 Muse 可执行服务 9

表 1： Jev 通过结构化决策与并行输出降低自动化调用开销 4
表 2： 省钱、工作与行政事务占公开 Muse 案例的前三位 7
表 3： 合作网络让 Muse 从信息获取进一步走向交易和预订 9

下载日志已记录，仅供内部参考

1 Jev：低成本决策模型探索 Agent 新分工

Jev 面向软件自动化中的快速结构化决策。9 月 15 日，前 OpenAI 研究者 Diogo Almeida 创办的 TypeSafe AI 发布首款 System One 模型 Jev，并开放早期访问。Jev 采用面向校准决策的强化学习 RLCD；开发者提供待判断的上下文与问题，模型返回可被代码直接使用的结果，适合嵌入已有工作流。

Jev 可将高频判断从长文本生成中拆分，有助于减少调用开销。Jev 提供 Choice、Score 和 Noul 三类输出，分别对应选项选择、按量表评分及真假判断概率，均不生成自由文本。选择与评分可返回概率分布和置信度；同一调用中的多个问题会基于相同上下文独立、并行评估，如一封邮件可以同时被判断紧急程度、归属类别和处理路径，系统再按结果调用不同工具。

图1： Jev 以三类结构化输出承接高频决策

Question type	Goal	Returns
Choice	Choose an option from a list	choice, probabilities, confidence
Score	Score the state on a rubric	score, probabilities, confidence
Noul	Is this statement true?	noul (0–1)

资料来源：TypeSafe 文档，浙商证券研究所

Jev 采用面向校准决策的强化学习（RLCD），训练模型直接返回结构化决策及概率，并使预测概率与实际发生频率相匹配，使软件能够据此自动处理或转交复核。Jev 直接输出程序可用的判断结果，并可在一次调用中同时完成多个独立判断。例如，对同一封邮件同时判断紧急程度、内容类别和处理路径，减少逐字生成回答及后续提取结果的开销，从而降低处理成本与等待时间。训练目标与推理方式围绕决策任务共同优化，为高频分类、评分和路由提供了低成本基础。

表1： Jev 通过结构化决策与并行输出降低自动化调用开销

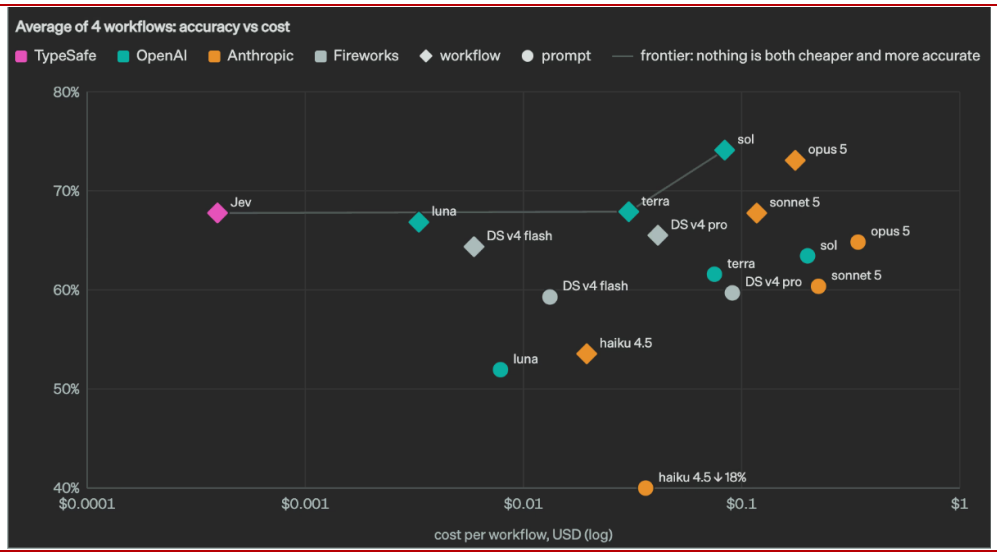
比较维度	通用大语言模型	System One 模型：Jev
训练方法	人类反馈强化学习（RLHF）、可验证奖励强化学习（RLVR）	面向校准决策的强化学习（RLCD）
优化目标	符合人类偏好的回答，或能够验证正确性的结果	结构化决策及与实际发生频率相匹配的概率
输入形式	以文本和连续对话为主	接收文本等非结构化数据，强调对程序状态的判断
输出形式	生成文本，接入软件通常需要解析和校验	直接返回预先定义的结构化结果及概率信息
输出方式	自回归生成，逐个 Token 依次输出	并行输出，一次调用返回多个决策结果
调用价格	输入价格约 0.20—10 美元/百万 Token，输出价格约为输入的 5 倍	输入价格为 0.042 美元/百万 Token，输出免费
响应时间	前沿模型端到端响应时间为 3—329 秒	端到端响应时间为 70—500 毫秒
不确定性表达	自报置信度可能存在过度自信或不一致	通过概率校准，使不确定性可用于软件分流
典型应用	对话、内容生成、编程及复杂推理	分类、路由、评分、校验及实时自动化工作流

资料来源：TypeSafe，浙商证券研究所

下载日志已记录，仅供内部参考

官方工作流测试显示显著成本优势。Jev 输入价格为 0.042 美元/百万 Token，输出免费。TypeSafe 的四项工作流平均表现图显示，Jev 位于准确率与成本的帕累托前沿，以更低成本达到接近 GPT-5.6 Terra 的表现；**公司同时披露，在所选工作流测试中，速度约为对照方案的 193.6 倍、成本约为其 1/444.6**。对需要反复判断、检查和分流的 Agent，单次决策变便宜，可使过去不值得自动化的细小环节具备部署价值。

图2： Jev 在官方工作流测试中以较低成本实现接近通用模型的准确率

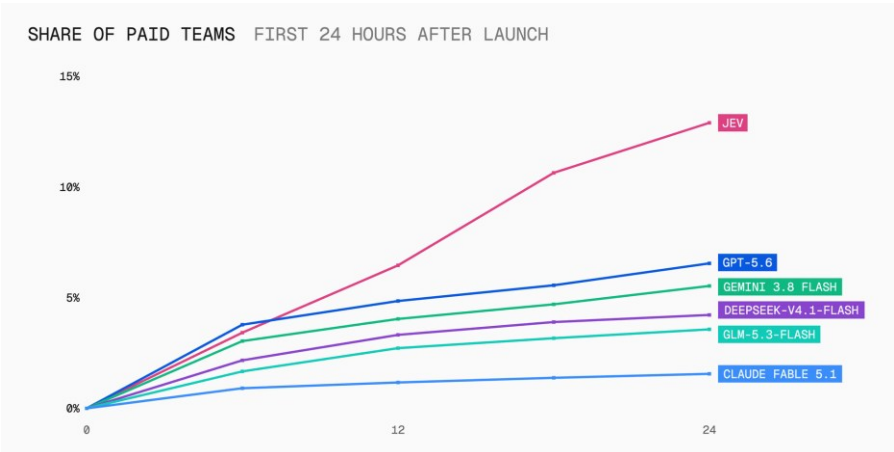


资料来源：TypeSafe AI，浙商证券研究所

置信度为自动执行与人工复核提供分流依据。高置信度且符合业务规则的结果可自动执行，不确定结果则补充信息、调用更强模型或转人工复核。开发者可按任务风险设定不同阈值，例如普通内容归类允许更高自动化比例，涉及付款或对外承诺则保留确认。因此，Jev 可作为 Agent 中的决策组件，与复杂推理和内容生成模型协同。

平台接入推动开发者快速采用。Vercel 于 9 月 18 日披露，Jev 上线 AI Gateway 后 24 小时内，已有近 13%的付费 AI Gateway 团队使用，成为该平台历史上采用速度最快的模型，首日采用比例超过此前模型发布的两倍。模型沿用现有网关接入与结算方式，有利于开发者将试验放入原有应用。此外，**Netlify** 已将 Jev 接入 AI Gateway，调用费用计入平台额度；**OpenRouter** 列出 Jev 模型并提供调用入口，**Cloudflare** 也已提供了接入。

图3： Jev 上线首日被近 13%的付费 AI Gateway 团队采用



资料来源：Vercel，浙商证券研究所

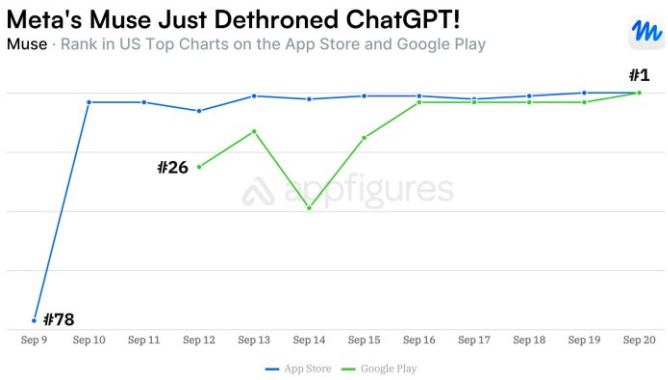
下载日志已记录，仅供内部参考

2 Muse：Meta 自研模型驱动，个人助理开始持续办事

Meta 于 9 月 8 日发布个人 AI Agent Muse，以自研 Muse Spark 模型提供支持。Muse 在专用云端虚拟机中运行，配备浏览器和个人数据空间，可通过应用或 WhatsApp 接收任务，关闭应用后仍能继续推进。其产品形态将逐轮问答扩展为目标委托：用户交代要完成的事务，Agent 规划步骤、调用工具，在需要确认时再通知用户。自研模型与 Agent 产品协同迭代，有利于围绕实际任务持续优化能力和成本；个人助理的竞争力也将由模型能力、工具执行与产品体验共同决定。

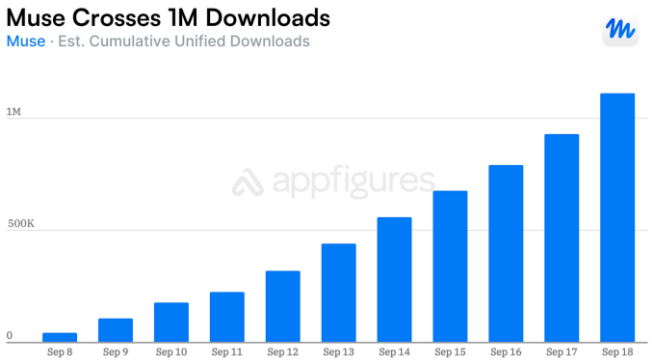
Appfigures 的排名图显示，截至 9 月 20 日，Muse 已同时位居美国 App Store 与 Google Play 免费应用总榜第一。其下载估算显示，9 月 8 日至 18 日累计下载约 110 万，主要来自美国与 App Store。另据数据分析机构 Sensor Tower 统计，Muse 于 9 月 8 日正式上线，上线仅 5 天下载量便突破 73 万次；上线 13 天（截至 9 月 21 日），全网累计下载量已超 250 万次，其中 iOS 端约 150 万次、安卓端约 110 万次。

图4： Muse 已登顶美国两大应用商店免费总榜



资料来源：Appfigures，浙商证券研究所

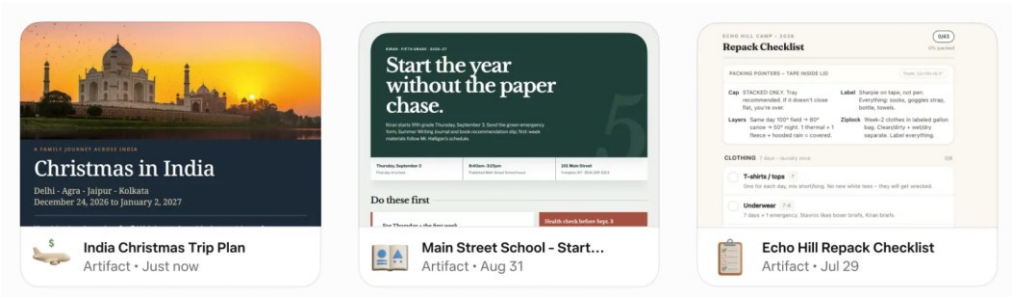
图5： Muse 上线初期累计下载快速增长



资料来源：Appfigures，浙商证券研究所

Muse 已有能力覆盖信息整理、任务执行与结果交付。Muse 的文件系统、终端与浏览器使其可以编写任务所需的工具，搜索网页、填写表单并处理预订和购买；输出可以是文档、网页、支出跟踪表或交互式仪表盘。Meta 披露的家庭场景中，Muse 可从学校邮件与网站整理开学事项，将日期放入家庭日历，准备采购清单并安排聚餐。这类场景的价值来自跨**信息源衔接**：用户交代一次目标，Agent 串联信息获取、整理与后续动作。

图6： Muse 可将跨应用任务交付为文档、工具和仪表盘

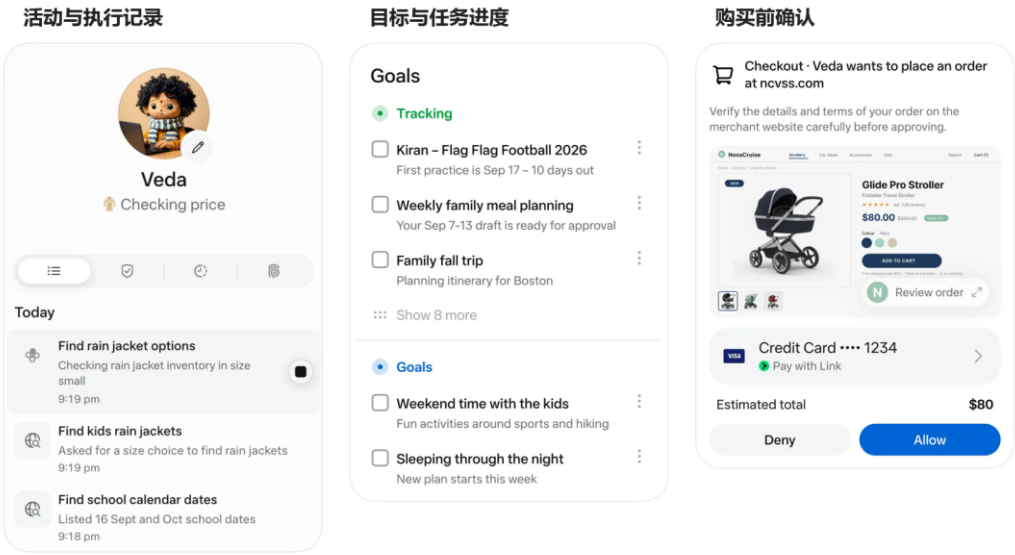


资料来源：Meta 官方产品示例，浙商证券研究所

下载日志已记录，仅供内部参考

记忆、目标与确认界面构成长期助理体验。 Muse 的目标页记录长期任务及推进计划，活动记录展示正在进行和已经完成的工作；用户可查看、编辑记忆文件，并管理已授予的权限。购买等关键操作通过明确的确认卡片处理。记忆减少重复解释，目标管理避免任务丢失，操作记录和确认机制使执行过程可见。上述持续使用环节，将直接影响用户的委托深度。

图7： 目标管理、活动记录和确认卡片支撑长期委托



资料来源：Meta 官方产品示例，浙商证券研究所

省钱、工作与生活事务率先形成需求。 研究平台 FUNDA 于 9 月 21 日对收集的 654 个公开 Muse 使用案例进行分类分析，发现节省开支、退款及个人财务共 128 例，占 19.6%；工作与小企业 73 例，占 11.2%；行政事务 69 例，占 10.6%；购物、技术与 Agent 设置、旅行分别占 9.5%、8.3%、8.1%。我们认为，账单沟通、资料整理和预订等耗时事务，更容易让用户直接感受到 Agent 的价值，也为付费提供了可量化的省钱、省时依据。

表2： 省钱、工作与行政事务占公开 Muse 案例的前三位

排名	主要类别	案例数	占比	典型任务
1	节省开支／退款／个人财务	128	19.6%	协商账单费用、更换保险、追回无人认领的资金、申请退款
2	工作与小企业经营	73	11.2%	重建网站、修复广告账户问题、寻找潜在客户、准备纳税申报材料
3	行政事务	69	10.6%	清理收件箱、处理文档、提交市政投诉、申请费用报销
4	购物	62	9.5%	寻找二手商品、下单、查询库存、办理退货
5	技术与智能体配置	54	8.3%	连接 Plaid／Robinhood、添加服务集成、在 Mac 上进行语音输入
6	旅行	53	8.1%	预订机票和酒店、取消行程、监测降价
7	个人事务管理	34	5.2%	管理日程、习惯及个人事务
8	健康	29	4.4%	预约或改期医疗服务、索取费用明细、寻找免费筛查
8	餐饮	29	4.4%	电话订外卖、预订餐厅、制定餐饮预算
10	家庭	28	4.3%	联系学校、安排儿童活动、申报受扶养人照护费用

资料来源：FUNDA 公开调查，浙商证券研究所

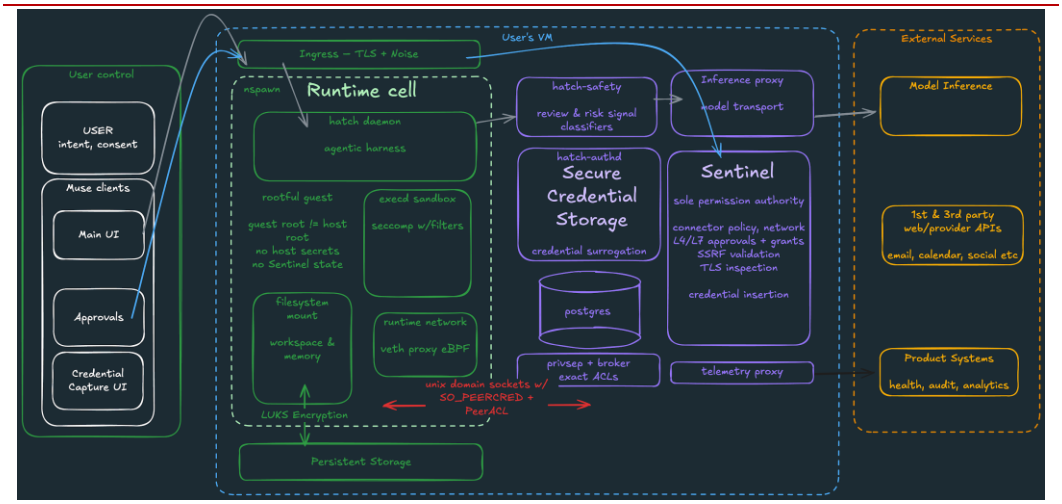
下载日志已记录，仅供内部参考

3 安全：账户授权与执行防护构成个人 Agent 的基础能力

Agent 安全需求正在从内容过滤延伸至身份、权限和运行过程。个人 Agent 会接触账户、数据与交易，安全需求随之扩展到谁在操作、能调用什么、是否超出授权以及事后能否追溯。我们认为，身份治理、运行时防护和审计有望形成新增需求，拥有相关产品和客户基础的安全公司具备拓展空间。

Muse 通过凭据隔离、独立审批与敏感信息过滤降低账户风险。 密码及 OAuth 凭据保存在主 Agent 运行环境之外，主 Agent 使用替代标记；外部连接与操作由独立的 Sentinel 审核，获准后才在网络边界注入真实凭据。用户可按任务、时段等范围授权。邮件连接器还会过滤验证码、密码重置链接等敏感内容，降低邮箱权限被进一步用于接管其他账户的风险。

图8： Muse 凭据隔离与独立审批将 Agent 操作限制在授权范围内



资料来源：Meta，浙商证券研究所

交易确认将自动执行限制在用户批准的范围内。Muse 每笔购买均需用户确认；Stripe Link 可提供限定商家、金额和有效期的一次性虚拟卡，使 Agent 能够执行获批交易，而不持有可任意使用的支付凭据。Meta 还计划在年内引入 Confidential VM，以强化云端运行环境的数据隔离。安全机制越能清晰限定授权范围，用户越有可能将账户连接、购物和长期任务交给 Agent。

4 生态：互联网服务持续扩展，AI 眼镜打开新的交互入口

社交入口有助于触达用户，服务连接决定任务能否落地，二者结合可将使用流量转化为持续委托。 Meta 的互联网生态同时提供用户入口和任务上下文。Muse 支持 WhatsApp 交互，并可连接 Instagram、Facebook 等服务；外部合作进一步补齐商品、支付、账户数据和预订能力。

Shopify 与 Stripe 推动购物能力向交易执行延伸。 Shopify 在 Muse 发布当日宣布，旗下商家已接入，Muse 可通过 Shopify Catalog 搜索匹配商品，并在应用内完成结账。Stripe 同日披露，Muse 可在超过 100 万家接受 Link 的商家使用用户保存的支付方式结账；面对其他商家，Link 可生成限定于获批交易的一次性虚拟卡，每笔购买由用户确认总额。

下载日志已记录，仅供内部参考

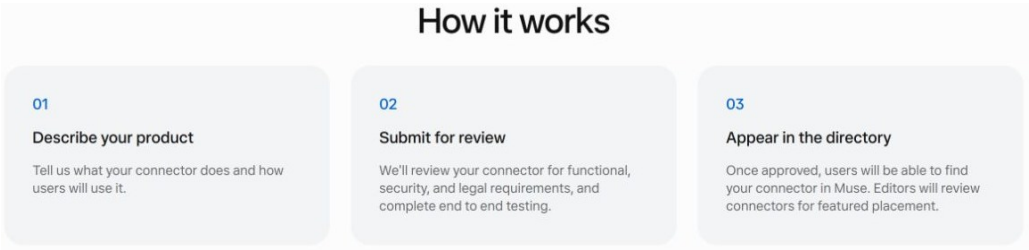
表3： 合作网络让 Muse 从信息获取进一步走向交易和预订

生态伙伴	已披露的接入能力	执行环节
Shopify	通过 Catalog 搜索商品，并在 Muse 内结账	商品与交易
Stripe / Link	覆盖超 100 万家 Link 商家；支持限定交易的一次性虚拟卡	支付与交易授权
Plaid	连接用户授权的余额、交易与持仓数据	财务信息
Duffel	搜索、预订及管理行程；平台覆盖超 500 家航空公司	旅行预订
Ticketmaster	发现活动与查询票务，最终购买在 Ticketmaster 完成	票务发现与导流

资料来源： Shopify，Stripe，Plaid，Duffel，Ticketmaster，浙商证券研究所

连接器平台将一次性集成扩展为持续生态供给。 Meta 已开放 Muse Connector Platform：开发者介绍产品和用途，提交功能、安全及法律要求审核和端到端测试，通过后进入 Muse 连接器目录；平台同时与 Stripe 合作支持 Link 收款。随着接入服务增多，个人 Agent 有望成为新的任务分发入口。

图9： 连接器审核与目录分发有助于扩大 Muse 可执行服务



资料来源：Meta，浙商证券研究所

AI 眼镜有望将个人 Agent 入口延伸至现实场景。Meta 与 EssilorLuxottica 推出的 Meta Glasses 已采用由 Muse Spark 支持的 Meta AI，Muse 发布公告称后续进入 AI 眼镜。眼镜的摄像头、语音和随身佩戴形态，有望降低用户交代任务的成本，将“看到、想到”更快转为可执行请求。我们认为，互联网服务解决“能够做什么”，智能终端则有望提升“何时被调用、使用是否自然”，共同扩大个人 Agent 的应用范围。

5 模型分工、服务生态与 AI 安全共同推进落地

Jev 与 Muse 分别从决策成本和执行能力两端，推动 Agent 扩大应用范围。 Jev 以面向校准决策的强化学习（RLCD）探索新的训练范式，通过结构化输出与并行判断，为 AI 办公、编程及安全类智能体中的分类、评分和路由环节提供更快、低成本的决策支持，并有望与承担复杂推理和内容生成的通用模型形成分工。Muse 则将 Meta 自研模型、长期记忆、账户授权与外部服务连接起来，使个人助理从提供建议进一步走向持续办理事务。我们认为，Jev 类决策模型带来了新的训练范式，将为 AI 办公、coding、AI 安全等智能体提供更快决策速度以及更低成本，推动用户为节省时间、降低费用和完成任务付费，**可关注 AI Agent 领域公司金山办公、道通科技、中控技术、金蝶国际等。**

下载日志已记录，仅供内部参考

个人 Agent 有望成为连接用户需求、互联网服务与智能终端的新入口，并带动 AI 安全需求扩展。 自有入口有助于降低获客成本、增加用户触达与任务发起频次，商品、支付和预订等服务接入使用户需求能够落地执行，眼镜等终端则通过随身感知与自然交互拓展使用场景。随着 Agent 获得更多账户权限并参与实际操作，身份认证、最小权限、运行防护与审计追溯将成为深化用户委托的基础能力。我们认为，个人 Agent 的普及有望拉动可信运行环境与安全防护需求，为网络安全厂商带来新的业务空间，**可关注网宿科技、深信服、绿盟科技、三未信安等安全公司。**

6 风险提示

应用效果及商业化不及预期的风险。若实际任务表现不稳定、人工介入较多，可能影响用户留存与付费转化。

服务接入及安全风险。外部平台限制接入、隐私泄露或错误执行，可能削弱产品功能与用户信任。

盈利能力承压的风险。价格竞争、免费补贴及持续运行成本，可能压缩产品利润空间。

下载日志已记录，仅供内部参考

股票投资评级说明

以报告日后的 6 个月内，证券相对于沪深 300 指数的涨跌幅为标准，定义如下：
1.买 入 ： 相对于沪深 300 指数表现 + 20％以上；
2.增 持 ： 相对于沪深 300 指数表现 + 10％ ~ + 20％；
3.中 性 ： 相对于沪深 300 指数表现 - 10％ ~ + 10％之间波动；
4.减 持 ： 相对于沪深 300 指数表现 - 10％以下。

行业的投资评级：

以报告日后的 6 个月内，行业指数相对于沪深 300 指数的涨跌幅为标准，定义如下：
1.看 好 ： 行业指数相对于沪深 300 指数表现 + 10%以上；
2.中 性 ： 行业指数相对于沪深 300 指数表现 - 10% ~ + 10%以上；
3.看 淡 ： 行业指数相对于沪深 300 指数表现 - 10%以下。
我们在此提醒您，不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系，表示投资的相对比重。
建议：投资者买入或者卖出证券的决定取决于个人的实际情况，比如当前的持仓结构以及其他需要考虑的因素。投资者不应仅仅依靠投资评级来推断结论。

法律声明及风险提示

本报告由浙商证券股份有限公司（已具备中国证监会批复的证券投资咨询业务资格，经营许可证编号为：Z39833000）制作。本报告中的信息均来源于我们认为可靠的已公开资料，但浙商证券股份有限公司及其关联机构（以下统称“本公司”）对这些信息的真实性、准确性及完整性不作任何保证，也不保证所包含的信息和建议不发生任何变更。本公司没有将变更的信息和建议向报告所有接收者进行更新的义务。
本报告仅供本公司的客户作参考之用。本公司不会因接收人收到本报告而视其为本公司的当然客户。
本报告仅反映报告作者的出具日的观点和判断，在任何情况下，本报告中的信息或所表述的意见均不构成对任何人的投资建议，投资者应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求。对依据或者使用本报告所造成的一切后果，本公司及/或其关联人员均不承担任何法律责任。
本公司的交易人员以及其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。本公司没有将此意见及建议向报告所有接收者进行更新的义务。本公司的资产管理公司、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。
本报告版权均归本公司所有，未经本公司事先书面授权，任何机构或个人不得以任何形式复制、发布、传播本报告的全部或部分内容。经授权刊载、转发本报告或者摘要的，应当注明本报告发布人和发布日期，并提示使用本报告的风险。未经授权或未按要求刊载、转发本报告的，应当承担相应的法律责任。本公司将保留向其追究法律责任的权利。

浙商证券研究所

上海总部地址：杨高南路 729 号陆家嘴世纪金融广场 1 号楼 25 层
北京地址：北京市东城区朝阳门北大街 8 号富华大厦 E 座 4 层
深圳地址：广东省深圳市福田区广电金融中心 33 层
上海总部邮政编码：200127
上海总部电话：(8621) 80108518
上海总部传真：(8621) 80106010
浙商证券研究所：https://www.stocke.com.cn