由 Codex 自动生成

AI前沿发展日报 | 2026-04-18（Asia）

日期：2026-04-18（Asia/Shanghai）

覆盖窗口：重点核查 2026-04-11 至 2026-04-18 期间新增信息，并补充少量 20
上旬仍在持续影响产业判断的高信号更新

今日总览

4 月 18 日最值得注意的，不是某一个模型分数再创新高，而是 AI 正在从“回答机器”加速变成“工作流操作层”。OpenAI 把 Codex 扩展到更多真实工具和自动化流程，Anthropic 把 Opus 4.7 的可靠性与高风险访问控制一起推进，说明企业采购开始更看重能否在权限、审计和长任务里稳定运行。

另一条主线是“可部署性”正在取代“纯参数想象力”。Google 把 Gemma 4 接进 Android 的 AI Core 开发者预览，Hugging Face 则推动 safetensors 进入 on，开放模型生态开始同时补齐端侧落地和供应链标准。

第三条主线来自产业现场。Microsoft 与 Stellantis 把 AI、工程协同、车联网和全打包成五年合作，说明大企业真正买单的，不是一个聊天入口，而是能把研发、制造、售后和客户体验一起改造的交付体系。

短期看，企业预算会继续向“能接权限、能本地跑、能进生产”的产品集中。中期看，模型公司的差异化将越来越来自工作流入口、部署形态、行业模板和治理能力，而不只是模型本身。

今日三条结论

1. AI agent 的下一阶段竞争，核心不再是“会不会回答”，而是“能不能跨工具、跨权限、跨时长地把任务真的做完”。
2. 开放模型与端侧部署已经从备选方案变成正式战略选项，硬件适配、模型封装标准和移动端分发都在进入主战场。
3. 对中国企业来说，最现实的机会不是追逐每一轮 frontier 发布，而是抢先做车厂、制造、客服、内容和移动终端里的本地化工作流重构。

今日 Top 5 大事件

1. OpenAI 把 Codex 推向“几乎所有工作”，coding agent 整的软件流程

发生了什么：OpenAI 在 2026-04-16 发布《Codex for （almost） e
ex 从代码补全工具继续扩展到更完整的开发工作流；OpenAI 在 2026-04-08 的企业更
中还披露，Codex 周活用户已达 300 万，OpenAI 当前约 40% 收入来自企业。

关键信息：新版 Codex 开始更强调跨应用操作、记忆、图像读取、终端与浏览器环境协同
，以及更长任务链条的自动化。企业更新则显示，OpenAI 每分钟已处理约 15 亿 tokens
，并把 Frontier、super assistant 与 Codex 并列为商业化重点，说明
正在从单点产品转向平台能力。

为什么重要：这意味着软件研发侧的 AI 竞争，已经从“谁能写出更像样的一段代码”升
级为“谁能接住产品、设计、文档、测试、排障和部署前后的一整串任务”。真正有价值的
不是一次生成，而是多步骤执行能力。

对产业 / 企业的启发：企业如果还把 coding assistant 当作 IDE 插件采购，很
估新一轮替代范围。更现实的采购标准会变成：能否接入 Jira、Git、设计稿、数据库、
终端、文档与内部权限体系；能否留下审计日志；能否支持异步长任务。

可信来源：OpenAI｜Codex for （almost） everything （https:
ex-for-almost-everything/） ｜ OpenAI｜The next phas
//openai.com/index/next-phase-of-enterprise-ai/）

2. Anthropic 正式发布 Claude Opus 4.7，把“更可靠的强
高风险访问”一起推进

发生了什么：Anthropic 于 2026-04-16 发布 Claude Opus 4.7 （
ic.com/news/claude-opus-4-7），并同步在 Claude、Anthropi
、Google Vertex AI 与 Microsoft Azure AI Foundry 上提

关键信息：Anthropic 将 Opus 4.7 定位为其最强公开模型，强调在复杂编码、视觉理解
、长文档与真实任务可靠性上的提升，同时保留与 4.6 相同定价。官方还明确把更高风险
网络安全能力继续放在验证门槛之后，通过专门的验证计划和防滥用控制分层开放。

为什么重要：frontier 模型正在形成新的发布范式。公开层面卖“更可靠的高能力模型”
，高风险层面卖“经过验证才能拿到的扩展权限”。这比单纯比跑分更接近真实商业化，也
更符合政企与关键行业的采购逻辑。

对产业 / 企业的启发：企业接下来在评估模型时，不能只看 benchmark 和单次回答质量
，更要看长任务稳定性、自检能力、风险分层与合规访问机制。做 agent 产品的团队，也
会更需要把审批流、日志、权限和回滚设计进产品底层。

可信来源：Anthropic｜Introducing Claude Opus 4.7 （https
ews/claude-opus-4-7） ｜ OpenAI｜Scaling trusted acc
ps://openai.com/index/scaling-trusted-access-for-

３． Ｍｉｃｒｏｓｏｆｔ 与 Ｓｔｅｌｌａｎｔｉｓ 签下五年合作，汽车行业开始把级改造工程

发生了什么：Ｍｉｃｒｏｓｏｆｔ 于 ２０２６-０４-１６ 宣布与 Ｓｔｅｌｌａｎｔｉｓ 扩大战略合来五年共同推进超过 １００ 个 ＡＩ、工程与数字化项目，覆盖客户体验、车辆软件、工程协同和网络安全。

关键信息：官方披露的合作范围包括产品开发、制造流程、客户服务、车内与车外数字体验，以及更广泛的安全与软件工程协作。重点不是上线一个 ＡＩ 助手，而是把模型能力嵌进车企从设计到售后的多条核心链路。

为什么重要：这说明传统大型产业客户已经不把 ＡＩ 视为创新部门的试点，而是开始按 ＥＲＰ、工业软件和云迁移那种规模去签长期合同。ＡＩ 项目正在进入真正的 ＣＡＰＥＸ/ＯＰＥＸ 决策层。

对产业 / 企业的启发：中国汽车、制造、供应链和工业软件团队需要警惕一个现实变化：未来订单不一定属于“模型最强”的厂商，而更可能属于“能理解复杂业务链条、能落地安全治理、能签长期交付合同”的集成型方案商。

可信来源：Ｍｉｃｒｏｓｏｆｔ｜Ｓｔｅｌｌａｎｔｉｓ ａｃｃｅｌｅｒａｔｅｓ ＡＩ-ｌｅｄ ｓｔｒａｆｏｒｍａｔｉｏｎ ｔｈｒｏｕｇｈ ｓｔｒａｔｅｇｉｃ ｃｏｌｌａｂｏｒａｔｉｏｎ ｗｉｔｈ Ｍｉｅｘｐｅｒｉｅｎｃｅｓ（ｈｔｔｐｓ://ｎｅｗｓ.ｍｉｃｒｏｓｏｆｔ.ｃｏｍ/ｓｏｕｒｃｅ/２０ａｔｅｓ-ａｉ-ｌｅｄ-ｓｔｒａｔｅｇｙ-ａｎｄ-ｄｉｇｉｔａｌ-ｔｒａｎｓｆｏｒｍａｔｉｏｎ-ｔｉｏｎ-ｗｉｔｈ-ｍｉｃｒｏｓｏｆｔ-ｔｏ-ｅｎｈａｎｃｅ-ｃｕｓｔｏｍｅｒ-ｅｘｐｅｒｉｅｎｃｅ

４． Ｇｏｏｇｌｅ 把 Ｇｅｍｍａ ４ 接入 Ａｎｄｒｏｉｄ ＡＩ Ｃｏｒｅ，开放动端原生入口

发生了什么：Ｇｏｏｇｌｅ 在 ２０２６-０４-１６ 发布 Ａｎｄｒｏｉｄ ＡＩ Ｃｏｒｅ Ｄｅｖｅ布把 Ｇｅｍｍａ ４ 带到设备侧推理体系中；此前 Ｇｏｏｇｌｅ 已在 ２０２６-０４-０２ 发布型家族。

关键信息：Ａｎｄｒｏｉｄ ＡＩ Ｃｏｒｅ 让开发者能够在兼容设备上调用本地 ＡＩ 能力，目标是把低延迟、离线可用、隐私友好的模型体验做成系统级能力。Ｇｅｍｍａ ４ 则延续开放许可与多规格路线，覆盖从更轻量部署到更强推理的不同场景。

为什么重要：移动端 ＡＩ 的竞争正在从“谁先有 ａｐｐ”转向“谁先占据系统能力层”。一旦开放模型能稳定进入 Ａｎｄｒｏｉｄ 原生栈，应用开发者就不必把所有体验都建立在云端 ＡＰＩ 上，成本、时延和隐私边界都会被重写。

对产业 / 企业的启发：对中国应用开发者、硬件厂商和内容平台来说，端侧模型将带来新的产品机会，包括离线客服、拍照理解、设备内工作流助手、教育工具和本地内容处理。下一轮价值不会只在模型公司，也会在终端集成、模型裁剪、蒸馏和芯片适配层释放。

可信来源：Android Developers Blog｜Android AI Core Deve
/android-developers.googleblog.com/2026/04/AI-Co
Google｜Gemma 4: Byte for byte, the most capable o
ogle/innovation-and-ai/technology/developers-tool

5. safetensors 进入 PyTorch Foundation，开源模型安全与标准治理”

发生了什么：Hugging Face 于 2026-04-17 宣布，safetensors 已正
undation 项目组合。

关键信息：safetensors 过去几年已经成为模型权重分发的重要格式，核心价值是避免传统序列化方式里的任意代码执行风险，同时提升权重加载效率。进入 PyTorch Foundation 后，它从单一社区项目升级为更广泛生态共同维护的基础设施标准。

为什么重要：开源模型这条线如今不缺模型本身，真正稀缺的是可信分发、兼容标准和可持续治理。谁能控制模型如何被安全封装、交换和加载，谁就更接近未来开源 AI 的“底层协议层”。

对产业 / 企业的启发：本地部署、私有化交付和多模型编排会越来越依赖这类底层标准。对中国企业而言，这直接关系到模型仓库管理、内部镜像、合规审计和供应链安全，不是一个只属于开发者社区的“小工具”话题。

可信来源：Hugging Face｜Safetensors joins the PyTorch F
ging Face.co/blog/safetensors-joins-pytorch-found
｜safetensors project page（https://pytorch.org/pr

商业与应用解读

对大模型公司来说，4 月中旬这一轮变化说明商业化层级继续上移。OpenAI 与 Anthropic 不再只是比“最新模型”，而是在比谁更能进入真实工作流、谁更能管理高权限场景、谁更能把长任务做稳。未来企业付费意愿最强的，不会是一次性问答，而是能长期嵌进研发、运营、文档、安全和审批流的 agent 层。

对 agent / coding / workflow automation 赛道，最值得关注的是 Codex 的方向说明 coding agent 正向“软件团队操作系统”演进；Anthropic 明强模型必须与分级访问和风控一起卖；汽车行业的大单则证明，真正的大合同来自把 AI 放进复杂组织流程，而不是做一个 standalone demo。这个赛道接下来最值钱的能力，是工具连接、状态管理、异步执行、权限治理和回滚。

对中国企业与内容服务场景，有三条更现实的落地方向。第一，端侧与私有化部署会继续升温，尤其适合车机、门店终端、客服、教育和企业知识库。第二，汽车、制造和供应链行业会更快接受“AI 不是项目，而是长期改造工程”的采购逻辑，本地集成商与行业 SaaS 会

有更大机会。第三，模型安全与分发标准会成为隐性门槛，谁能把模型封装、镜像管理、审计和治理做好，谁更容易拿到政企与大客户订单。

还有一个值得单独跟踪的信号是垂直化。OpenAI 在 2026-04-16 发布 GPT-Rosa（https://openai.com/index/introducing-gpt-rosalind/），面向生命科学与药物研发场景。它未进入今天 Top 5，是因为短期商业外溢速度还不如工作流与端侧部署明确，但它提醒市场：通用模型平台的下一轮价值，很可能来自少数高价值行业模型，而不是对所有行业一视同仁。

X 平台高信号观点

1. @OpenAIDevs：Codex 的命题已经不是“帮你写代码”，而是“帮你更多工作”

类型：已验证事实 + 趋势信号

验证状态：Codex 扩展方向已被 OpenAI 官方产品页验证；“从代码走向工作”是基于其新能力边界的趋势判断。

一句话判断：coding agent 的竞争边界正在从开发环节外溢到整个知识工作流。

来源：OpenAI Developers on X（https://x.com/OpenAIDe 5282）｜OpenAI｜Codex for（almost）everything（htt x-for-almost-everything/）

2. @PyTorch：Gemma 4 的关键不是继续堆大，而是把 intelli yte 做到更有部署价值

类型：趋势信号

验证状态：“intelligence per byte”来自 PyTorch 的公开表述；Gemm格和端侧定位已被 Google 官方页面验证。

一句话判断：开放模型下一轮竞争会更像系统工程，而不是单纯参数竞赛。

来源：PyTorch on X（https://x.com/PyTorch/status/204 le｜Gemma 4（https://blog.google/innovation-and-ai s/gemma-4/）

3. @tanayj：高风险前沿模型的商业化，不会默认走向公开发布

类型：观点 + 已验证事实

验证状态：关于 Claude Mythos Preview 内部时间点、网络安全能力与未公开发布原

关键信息，可被 Anthropic 官方 system card 验证；“frontier 模型会
放”是趋势判断。

一句话判断：越强、越敏感的能力，越可能先进入验证名单和封闭预览，而不是直接进入通
用 API。

来源：Tanay Jaipuria on X（https://x.com/tanayj/stat
oto/1）｜Anthropic｜Model system cards（https://ww
ds）

前沿研究速递

1. Action Images：让机器人策略直接建立在“可解释动作图像”上

做了什么：这篇 2026-04-15 更新的论文提出 Action Images，把机器人动作编码
释的多视角动作图像，并把策略学习统一到视频生成框架里。

新在哪里：它不再把控制信号当作抽象 token，而是把动作直接投影到像素空间，让视频
骨干网络本身就能充当零样本策略，而不必额外再接一个独立 policy head。

潜在应用方向：机器人抓取、工业臂控制、仓储自动化、仿真训练、跨视角操作迁移。

一句话判断：如果机器人策略能直接继承视频模型的表示能力，具身智能的训练成本和迁移
效率都有机会被重写。

来源：arXiv｜Action Images: End-to-End Policy Learnin
neration（https://arxiv.org/abs/2604.06168）

2. VGA：把机器人通用控制从“视觉到语言”改写成“视觉到几何”

做了什么：这篇 2026-04-14 发布的论文提出 Vision-Geometry-Action
接基于 3D 世界表征生成动作，而不是依赖传统视觉语言或视频骨干。

新在哪里：作者认为机器人操控的本质是从视觉到几何的映射，因此把原生 3D 表征而不
是语言语义放到控制核心，并在真实世界零样本视角泛化上优于多种 VLA 基线。

潜在应用方向：精密抓取、装配、复杂操控、工业机器人、具身智能底座模型。

一句话判断：具身智能下一轮关键分歧，可能不在语言能力，而在 3D 几何表征是否足够
原生。

来源：arXiv｜Robotic Manipulation is Vision-to-Geomet
tarrow G$): Vision-Geometry Backbones over Langua
/arxiv.org/abs/2604.12908）