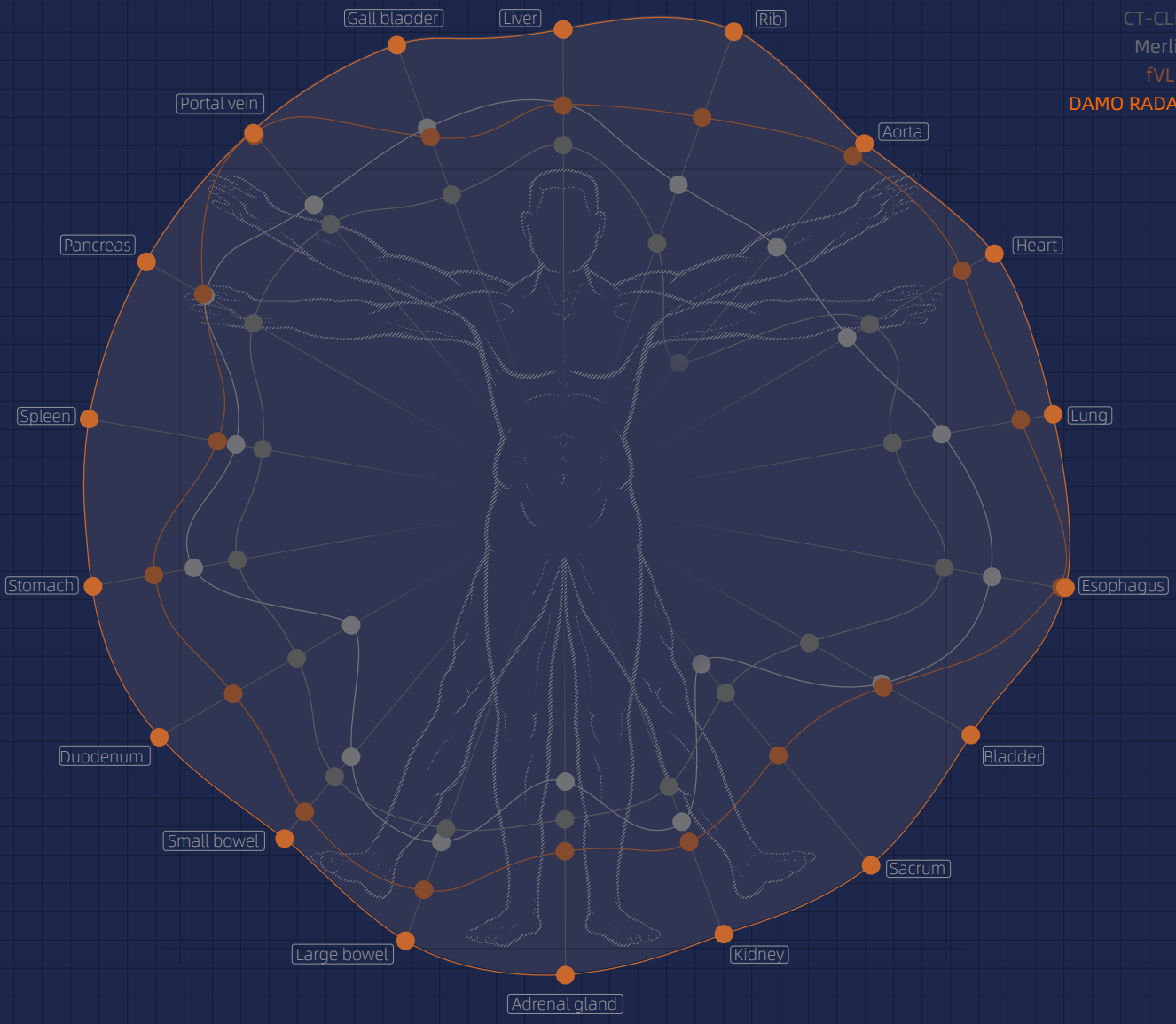
Alibaba



阿里巴巴
AI 向善行动报告 2026

AI for Good

封面说明：
阿里巴巴达摩院专家级通用医疗影像 AI 模型 DAMO RADAR 正式开源，已经覆盖 18 个器官和 146 种病症，成果在《科学》杂志正式发表。封面图选自论文。

目录

序言一 向善，站在智能体时代的起点 2

序言二 向善而行，就要向众而行 4

第一章 AI 与社会共同演进 6

当 AI 加速走入社会 7

技术与社会如何相互适应 7

构建 AI 可持续发展的条件 8

第二章 全栈 AI，协同向善 9

构筑 AI 安全的纵深防御 11

开源开放鼓励多方参与 13

绿色基础设施和算电协同 15

第三章 向善行动 2026 17

AI 赋能健康 19

AI 赋能教育 24

AI 赋能科研 28

AI 赋能绿色发展 31

AI 赋能经济机会 34

AI 赋能安全包容的社区 37

第四章 面向未来的思考 41

进入社会，AI 再次演进的新起点 42

求真与向善：面向未来的八个问题 43



向善，
站在智能体时代的起点



过去一年，AI 跨入新的发展阶段，从“回答问题”走向“执行任务”，智能体（Agent）能够自主规划、调用工具、协调多个子系统、完成连续决策，进入真实社会场景、逐渐成为能够“做事”的系统。我们一起见证了这一远超技术本身的转变，站上了智能体时代的起点。

AI 技术发展不是被动的无序演进，而是利益攸关方协同“校准”的过程。历史上，人类的技术发展总是面临类似的挑战：在技术发展早期，人们很难预测其社会影响，一旦技术成熟并深度嵌入社会，再想改变方向又代价高昂。在我看来，不能等所有答案都清楚了再行动，而要在不确定性中本着向善初心，持续行动、不断调整，始终确保技术发展方向符合人类福祉、创造社会价值。

当然，挑战真实而具体。在我看来，AI 发展的现实路径需要满足三个条件：社会可接受、环境可承载、经济可持续。

社会可接受，当智能体进入真实社会场景，其服务必须可靠可信，也要让人切实受益。在今年的报告中我们看到，千问推出志愿填报智能体，帮助高考生完成人生第一次大考之后的学校和专业选择；阿里巴巴国际站 Accio 帮助 1000 万中小企业开展跨境贸易业务；阿里健康“氢离子”开始辅助医生进行临床决策。当智能体开始辅助人做出决策，用户对 AI 的信任就不再是一个抽象概念，而是每一次具体交互中可以被验证、被追溯，并随着服务迭代不断积累的过程。

环境可承载，过去一年全社会对 AI 行业最集中的关切之一，是其能源和水资源消耗。2026 财年，阿里巴巴自建数据中心使用清洁电力的比例已达 73.6%，并在广东、江苏、四川实现 100% 覆盖。但当算力需求还在大幅增长，单位效率的提升并不必然带来绝对消耗的下降。如何在扩大 AI 能力的同时控制碳排放、降低对水资源的影响，是整个行业必须共同回答的问题。

经济可持续，关乎 AI 能否创造与投入相称的真实价值。AI 基础设施的资本投入强度前所未有，我们期待并相信，这些投入的价值绝不止于给各行各业提效，AI 技术将不断解决真正重要的社会问题、惠及更广泛的人群，并给全社会带来更长期的改变。

作为全球领先的 AI 公司，阿里巴巴一直高度重视 AI 本身的安全和伦理治理。2022 年底，我们成立科技伦理治理委员会，此后又发布了《阿里巴巴集团科技伦理治理准则》。当 AI 从对话框发展到智能体，安全的含义也随之扩展，不仅要关注模型说了什么，还要关注智能体做了什么。为此，阿里巴巴已构建了覆盖模型全生命周期的六层安全纵深防御体系。

同时，我们也坚定选择开源路线，让向善的力量可以触达更大范围的开发者生态。截至 2026 年 9 月，阿里巴巴累计开源 460 余款千问 Qwen 系列模型，衍生模型超过 30 万个，全球累计下载量突破 30 亿次，支持超过 200 种语言和方言。

序言

在 Hugging Face 上，全球开发者自主完成了数以万计的千问模型本地部署转换版本。由阿里创建并运营的魔搭社区，汇聚了 3300 万开发者和超过 25 万个开源模型。我们相信，一个由社区驱动的开放生态，比封闭系统更具韧性和创造力。

能否帮助解决真实的社会问题，是技术最大的试金石。每一个向善场景的复杂性和多样性，本身就是训练更强 AI 的稀缺资源。例如，从胰腺癌的 AI 识别技术出发，今年达摩院的通用医疗影像 AI 模型 DAMO RADAR 已覆盖 18 个器官、146 种病症。这显示出 AI 在真实社会场景中不断校准、迭代、创造更大社会价值的巨大潜力。

当然，新生事物总是与问题相伴，AI 也不可避免；有些问题全世界仍然没有找到成熟的答案，这是挑战，也是机会。这意味着我们依然处于技术革命的规则形成期，可以基于我们的思考和实践，去探寻一条通向人类福祉的 AI 发展之路。探索贯穿在我们每天的工作中，也体现在这份报告里 —— 我们解决了什么，正在做什么，还有哪些挑战。这是我们交出的一份阶段性答卷，也是面向同行和伙伴的一次坦诚对话。

方向是确定的，路要一步一步走，而每一步都需要更多同行者。

吴泽明

阿里巴巴合伙人
阿里巴巴集团首席技术官
阿里巴巴集团科技伦理治理委员会主席

向善而行，就要向众而行



如果说去年 AI 更像是实验室里的“重器”，那今年就更像是走入千行百业的“众器”。能对话、也能办事，AI 越来越多地进入医院、学校、社区、乡村，走进我们每个人工作和生活的日常。

看待 AI 发展，实验室里的技术参数提升当然重要，但更重要的是：能否可靠地为人所用、真实地让人受益。这意味着要真正实现 AI 向善，离不开一个人人可及、人人受益的社会系统。

高精尖的技术，不会自动进入普通人的生活；要让普通人用得上、用得好，我们需要为技术和社会的融合创造条件。要做到这一点，离不开社会各界的合力，而企业探索正是其中的重要组成部分。今年报告揭示了很多阿里在这方面的思考和实践。

首先，我们要依据社会效应来为 AI 应用设定优先赛道，优先探索技术服务最大多数群体福祉的可能性。我们去年的报告就提出：“将技术能力投向最能系统性产生影响的关键领域。”这一初衷没有变化，既然 AI 是人类社会迄今投入强度最大的技术创新，我们理应优先将其用到能最大化解决社会问题的关键领域。

比如，在供需矛盾突出的医疗领域，AI 的潜力不可限量。达摩院的“平扫 CT+AI”早筛技术覆盖面越来越广，从最初的单一癌症筛查，今年已经覆盖了五种消化系统癌症，并可通过腹部增强 CT，一次性识别超过 146 种病症。这种技术不仅能挽救生命，如广泛应用也能有助于缓解公共医疗资源压力。

又比如，在科研领域，AI 展现出了在病毒鉴定、材料发现、细胞研究等前沿领域的巨大应用前景，“AI 加速科研、科研反哺 AI”的递归式迭代正在形成。选择科研赛道，也恰恰是因为科学的每一次突破，成果都可能反哺到每一个人。

其次，我们要创造让更多人用得上 AI 的机会。AI 时代的可能性，不是 AI 的可能性，而是人的可能性，或者说是人加上 AI 之后的可能性。

例如，我们通过“少年云”项目普及 AI 教育应用，是希望实现乡村孩子在 AI 时代的发展可能性；我们通过循证医学大模型“氢离子”辅助诊断，是希望让中国数百万医生，尤其是边远地区医生也能有持续学习和提升诊疗质量的可能性；我们通过魔搭社区普及开源模型，是希望让更多开发者在普惠算力和开放生态中找到解决社会问题的可能性。

再次，我们要探索出一个技术能力和社会场景之间的协同机制，来帮助 AI 技术在越来越多的场景中去落地。今年，AI 行业有个新职业很火：前沿部署工程师（FDE）。我们在公益一线观察到的情况，印证了这种需求。

比如在乡村振兴领域，仅仅把技术能力送下乡是不够的，你需要理解基层的真实

序言

需求，帮助技术融入本地组织流程。阿里乡村特派员今年在浙江嵊泗，将复杂船票体系、航线规则和天气变化等本地经验，与技术能力整合为 AI 船票助手，解决了当地一大出行难题。类似的县域实践正在更多地区展开，这种协同机制是决定 AI 能否真正走入越来越多真实场景的关键。

这些一线实践提醒我们：AI 有能力，但把能力转化为结果，靠的不是技术本身，而是人。我在一次调研过程中听到一位县域医生的感叹：AI 早筛技术的确可以发现癌症风险，但医生能否准确核实，医院能否成功叫回病人，病人是否愿意配合复检，复检后能否正确转诊……只靠 AI 是远远不够的，大量现有和未来的工作还需要人的判断和参与。

AI 很强大，但 AI 不是万能的；在强大的 AI 和真实的社会问题之间，离不开人的努力，离不开把人凝聚在一起的制度和文化。如果我们想享受 AI 的红利，那我们就必须共同打造一个有爱、有韧性的社会。

这一年，我看到阿里人在为推动技术迭代发展勇攀高峰，我也看到阿里人在为实现技术落地不懈努力：比如报告中的“追星星的 AI”，就源于阿里自闭症儿童家长们的自发探索；阿里工程师还利用休息时间为杭州余杭区打造守护独居老人的 AI 安全方案，阿里技术小二为长江源的野生动物保护设计 AI 识别方案……

我们以前常说，有爱的人才能打造有温度的技术和产品，而这样的技术和产品才能定义一家好公司。在 AI 时代，不管技术如何发展，这样的初心没有变化。希望这份报告成为一块小小的路标，标记我们走过的路，也激励我们继续前行。

孙利军

阿里巴巴集团公益总裁

阿里巴巴公益基金会理事长

第一章
AI 与社会共同演进



当 AI 加速走入社会

“基础设施建设和资本投入推动技术能力形成，但技术能否真正产生社会价值，还取决于组织能否有效应用、制度是否建立，以及市场能否支持服务长期运行。

2026 年，AI 的发展迎来了一个清晰的拐点，开始从技术突破迈向规模应用。工业革命以来，机器动力逐步替代了人类大部分体力劳动；当机器智能开始规模化地分担思考与决策工作，技术对社会的影响或许才真正开始。

模型应用形态的改变是这一阶段的关键。不同于早期的对话交互，今年的智能体已经能够规划任务、调用工具并参与业务流程，推动 AI 应用从辅助问答走向复杂的任务执行，逐步进入千行百业和人们的日常工作与生活。与此同时，对芯片、云基础设施、模型、应用等的持续投入，则进一步扩大了机器智能的规模化供给，让更多组织和个人能够获得可用的 AI 工具与服务。

当 AI 加速进入社会，也带来更直接的现实问题。模型能力的提升并不自动带来可靠性的增加，激烈的行业竞争让安全对齐和风险评估面临更大挑战；数据中心的扩张加剧了能源和水资源的消耗，增加环境负担；就业替代和职业转型的担忧也受到更多关注。

技术能力的形成并不自然带来社会广泛受益，还需要相应的组织、制度和市场条件。基础设施建设和资本投入推动技术能力形成，但技术能否真正产生社会价值，还取决于组织能否有效应用、制度是否建立，以及市场能否支持服务长期运行。这两方面相辅相成、相互影响，妥善的制度安排和治理机制是技术长期可靠运行的重要条件。

在这一转化过程中，纯粹的市场逻辑难免存在盲区。一些重要的社会需求，由于受众支付能力有限、服务成本较高或回报周期较长，容易被市场机制忽视。要弥补这种失衡，既需要公共政策、科研支持和企业的长期投入，为这些领域保留探索空间；也依赖技术进步和规模化供给带来的降本，让原本难以覆盖的社会需求获得可行的解决路径。

因此，理解 2026 年的 AI 向善，不能只关注应用之后的福祉和风险问题，还需要将视线前移到技术研发、资源配置和场景设计的最前沿，让向善从一开始就成为技术发展的内在考量。

技术与社会如何相互适应

AI 从技术能力形成到大规模进入社会，需要经历一个适应和融合的过程。能否顺利完成这一过程，取决于两个方面：技术能否回应真实需求，以及社会是否具备使用、评价和约束技术的能力。

这意味着，技术的发展方向并不完全由技术本身决定。社会需求、制度选择和公众参与都会影响技术方向，决定创新优先回应哪些问题，以及哪些人能够受益。在一些场景中，改进组织流程和制度安排，比单纯提高某项技术指标更能改善实

际效果。

在这个过程中，关键不只是让 AI 变得更聪明，更重要的是让它更“合群”。AI 在具体场景中应用，需要在明确的规则和责任下运行。身份如何明确、权责如何厘清、边界如何界定、出现问题如何追责，都需要相应的制度性安排。缺少这些安排，技术能力越强，潜在影响和社会面临的系统风险也可能越大。

当前，各主要经济体都在推进产业部署的同时完善治理框架，力求让制度建设跟上技术应用的节奏。但 AI 技术迭代速度快，治理不是一次性的制度设计，而需要持续调整的能力。这需要多方协作：法律法规明确基本要求，标准和实施措施根据新的证据不断完善，企业在部署和运营中记录问题，专业机构提供独立评估，使用者和受影响群体提供反馈。

同时，个人和组织也需要逐步形成适应 AI 的能力。企业可能需要调整流程和岗位，教师需要利用 AI 支持学习的同时保留必要的思考和练习，医生使用 AI 辅助诊疗时仍需保持专业判断。这些调整越早开始，大规模应用时的成本和难度就越低。

技术系统和社会系统需要协同演进，才能让新技术惠及更多人。当 AI 能够顺利嵌入社会系统，而社会系统也具备吸收、适应、约束 AI 能力的时候，AI 才真正从一种前沿技术，转化为一种被信任、被共享、可持续创造公共价值的社会基础设施。

构建 AI 可持续发展的条件

“实现技术系统和社会系统的协同演进，需要以“社会可接受、环境可承载、经济可持续”三方面为基础，三者共同构成 AI 能平稳有序发展的“安全运行空间”。而只有具备这些条件，AI 向善发展的潜力才能得到充分释放。

实现技术系统和社会系统的协同演进，需要以“社会可接受、环境可承载、经济可持续”三方面为基础，三者共同构成 AI 能平稳有序发展的“安全运行空间”。而只有具备这些条件，AI 向善发展的潜力才能得到充分释放。

从社会维度来看，AI 的大规模部署离不开公众接受和监管认可。安全事故、系统不透明、对破坏性结果的担忧，都会压缩这个空间。社会能否接受 AI，取决于服务是否可靠可信，以及不同群体能否真正受益。保障安全、提高透明度，在出现问题后及时纠正并提供救济，是建立信任的基础。

环境承载力的考量则涉及能源、水资源和碳排放等多个方面。优化模型、芯片和运行调度能够提高资源利用效率，使用清洁能源有助于减排，但这些改善未必能抵消规模增长带来的总体影响。因此，环境约束需要从一开始就纳入 AI 的规划、建设和运营全过程，不能等到规模扩大后再作补救。

从经济维度看，AI 的投入强度前所未有。能否持续发展，取决于 AI 能否解决足够重要的问题，创造与投入相称的实际价值。同时，还需要形成支持长期运行的机制，让短期内商业回报有限的服务能够持续探索并最终发挥价值。

这三个维度相互影响，也可以相互促进，让技术发展与社会受益形成良性循环。维护和拓展这一“安全运行空间”，有赖于各方的共同努力，掌握更多技术和资源的机构尤其应承担与其影响力相称的责任。向善不是在业务之外附加的社会责任，而是贯穿技术研发、基础设施建设和产品服务全过程的内在要求。



第二章
全栈 AI，协同向善

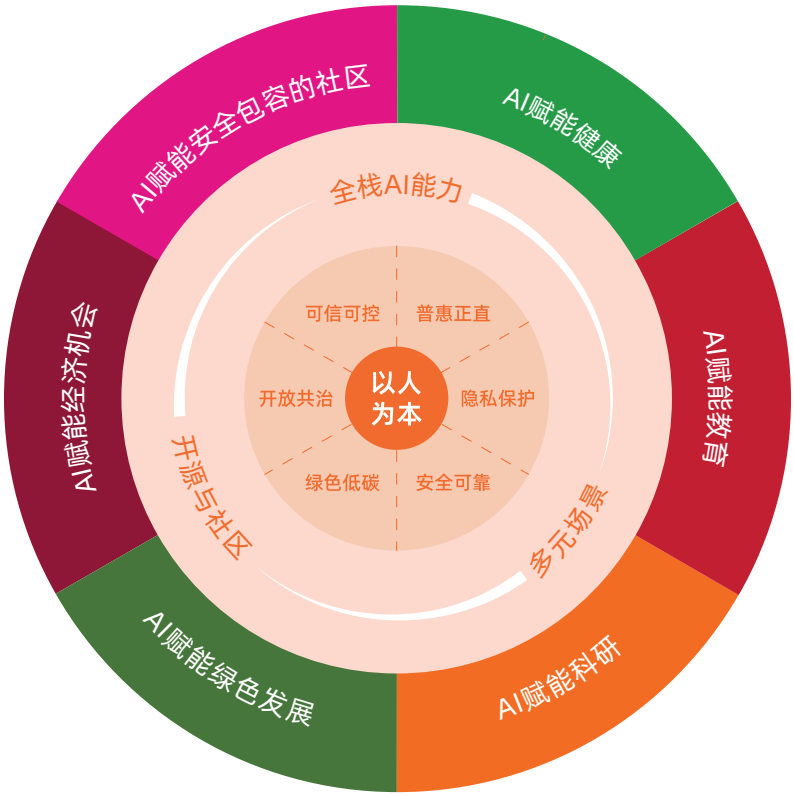
在去年的报告中，我们提出了阿里巴巴 AI 向善的行动框架：以“以人为本”为价值导向，通过六项原则指导技术发展，以普惠的基础设施赋能个体和组织，并在具体场景中探索解决社会问题。随着 AI 从内容生成走向任务执行，更深入地融入组织流程和日常生活，技术能力的协同、开发者与合作伙伴的参与，以及真实场景中的验证和反馈变得更加重要。

今年，我们结合阿里巴巴在技术、生态和业务场景中的积累，进一步完善框架，将全栈 AI 能力、开源与社区、真实场景与反馈明确为实践支撑。我们的实践方向也拓展为健康、教育、科研、绿色发展、经济机会、安全包容的社区六个方面，与多项联合国可持续发展目标相联系。框架的出发点始终不变：技术能力是手段，让人切实受益才是目的。

这一框架的完善，也基于阿里巴巴过去一年在技术建设与应用中的实践积累。我们通过模型、芯片与推理系统的协同优化，提升运行效率、降低使用成本；通过开放模型、工具和社区协作，支持更多开发者与合作伙伴参与创新；通过真实场景中的应用与反馈，持续改进技术和服务。我们也在研发、部署和运行环节完善安全措施，并探索通过资源调度和清洁电力利用，提高基础设施的资源与能源使用效率。

这些实践需要面对不同目标之间的权衡。全栈协同并不意味着性能、成本、安全和环境目标总能同时改善，但为统筹这些要求提供了条件。我们关注技术发展对人的实际影响，持续探索向善原则与业务实践的结合。下文将聚焦支撑 AI 向善应用的基础性工作，从安全建设、绿色基础设施和开源开放三个方面，介绍过去一年的实践与进展。

阿里巴巴 AI 向善行动框架



以人为本的 AI 向善行动框架

治理与披露进展

2021
成立集团可持续发展委员会；提出云计算碳中和战略与目标。

2022
成立集团科技伦理治理委员会；发布《人工智能治理与可持续发展实践白皮书》。

2023
科技伦理治理纳入集团可持续发展工作体系；发布《生成式人工智能治理与实践白皮书》。

2024
发布《阿里巴巴集团科技伦理治理准则》《阿里巴巴集团科技伦理审查管理规范》及《大模型技术发展及治理实践报告》。

2025
发布首份《阿里巴巴 AI 向善行动报告》。

2026
联合发布人工智能安全白皮书《守己利他：智能时代做负责任的技术》。

构筑 AI 安全的纵深防御

“当 AI 能力从“说话”延伸到“行动”，更需要安全防护依照“零信任架构”来组织。

安全可靠是阿里巴巴 AI 向善的六项指导原则之一。落实这一原则，阿里巴巴关于模型安全的基本理念是“守己”：不是保守地限制模型能力，而是在模型能力快速扩张的同时，让模型始终守住安全、合规、可信和可控的边界。

当 AI 从对话框发展到智能体，安全的含义也随之扩展，不仅关注模型说了什么，还要关注智能体做了什么。为此，阿里巴巴构建了覆盖模型全生命周期的六层安全纵深防御体系：评测牵引、内生安全、外层护栏、数字水印、云上安全、智能体安全。

第一层是评测牵引。阿里巴巴的评测体系，能够刻画模型的真实安全水位，既覆盖模型基座、推理过程和最终输出，也覆盖文本、图像、音频、视频和跨模态场景，还能进入真实业务场景。阿里建设了 S-Eval 等安全评测体系，并将常规评测、红队评测、攻击式评测、神经元知识探针结合起来。评测不仅能发现问题，还能把风险知识点转化为训练、对齐和护栏优化的依据。

第二层是内生安全，让安全成为大模型的本能。安全不仅是在外部护栏系统设置拦截，更重要的是通过模型架构、训练数据治理、对齐技术和任务设计，把安全能力编译进模型内部，更好地激发模型自身的安全意识和能力，仅在模型的安全缺陷或薄弱处，以及个别风险极为关键和结果敏感的场景，发挥护栏的阻断作用。

阿里发布的模型规约，把抽象的安全价值变成模型可以学习、遵守、评测和迭代的行为规则。它更接近模型行为的“宪法”：当用户请求、开发者政策和根本原则发生冲突时，规定了服从安全原则的高优先级，同时在安全边界内尽量满足用户的合理需求。

规约建立了“根准则 > 系统准则 > 开发者准则 > 用户偏好”的指令优先级结构，使模型在冲突场景下能够更可靠地服从高层级安全要求，同时保留开发者二次开发的可控性。

第三层是外层护栏。部署在模型推理服务链路外围，可以在用户输入进入模型前、模型输出到达用户前进行风险检测和智能干预。QwenGuard、PlugGuard、EDS 等

指令优先级的模型规约：层级体系将价值声明转化为工程级目标

层级	设置方	可覆盖性	设计意图
根	模型规范	不可被覆盖	跨文化共识底线
系统	模型提供方	仅可通过系统指令	司法管辖区合规与受监管用户群体
开发者	应用创建者	可通过开发者指令	差异化 AI 应用
用户	终端用户	可通过开发者或用户输入	个性化与一致性

来源：阿里巴巴集团安全部

原生护栏技术，尽量不改动原模型，就能提升复杂交互中的安全防御能力，降低误拒率，同时保持高安全性。尤其是流式原生护栏，可以在模型生成过程中利用中间隐藏状态提前感知风险，在有害内容真正输出前进行拦截。

第四层是数字水印与 AI 生成内容检测，是防止 AI 盗用和滥用的关键技术。“检测–水印–检验”的闭环体系，能先识别 AI 生成内容，再对其进行标记，包括显式标识、文件元数据隐式标识、内容隐式水印等方式，最后支持追溯和合规验证。该体系覆盖文本、图像、音频、视频等多模态内容，并已经在高德、闲鱼、淘宝、天猫等业务场景中落地。它说明模型安全已经延伸到生成内容的流通、识别、溯源和治理。

第五层是云上大模型服务安全。公共云不仅提供模型，还要为模型训练、推理部署、工具调用、服务编排、监控运维等全生命周期提供安全保障。在云上大模型服务中，我们强调纵深防御、零信任、最小权限，把身份验证、访问控制、运行时防护、数据生命周期保护和 DevSecOps（开发安全运维一体化）嵌入 AI 基础设施。

第六层是智能体安全。随着智能体进入业务流程，需要关注的风险从模型输出本身，扩展到智能体的感知可靠性、决策可解释性、权限控制和多智能体协同。为此，阿里巴巴为智能体建设了专属的权限体系和插件沙箱机制，在网络、执行、用户、权限和数据等层面实施隔离。通过全链路审计记录智能体的身份、操作类型和关键数据标识，并配合独立的智能体行为护栏，实现异常监测、意图偏离检测，以及遭受攻击或失控时的溯源和取证。

阿里的模型安全可以概括为：不是在模型之外临时加一道安全闸门，而是把安全写进模型、写进工程链路、写进云基础设施，也写进智能体进入真实世界的权限、身份、审计和责任结构之中。

智能体安全整体思路

三位一体，构建安全可信的智能体体系



来源：阿里巴巴集团安全部

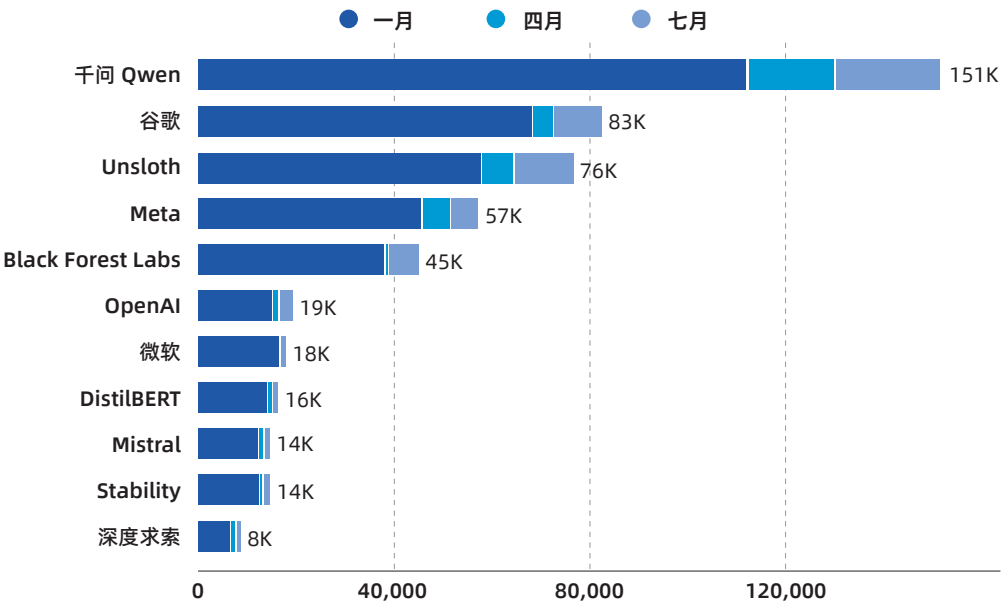
开源开放鼓励多方参与

仅靠个别企业的力量，AI 无法广泛进入社会。AI 进入千行百业的过程，需要研究者、开发者、企业、社会组织和公共机构等不同主体依托其行业知识和能力，在各自场景中构建具体的 AI 应用。

作为一家平台企业，阿里巴巴能做的，是通过多种方式将 AI 能力开放给不同群体：开源模型，面向有技术能力的开发者和研究机构；云平台与 API 服务，面向需要 AI 能力但缺乏自建条件的企业和组织；产品化应用，则将 AI 能力带给不需要技术背景的知识工作者和普通用户。

- **模型开源：降低 AI 技术门槛、加速 AI 技术的扩散和普惠。**截至 2026 年 9 月，阿里巴巴累计开源 460 余款千问 Qwen 系列模型，衍生模型超过 30 万个，全球累计下载量突破 30 亿次，并在阿拉伯语、日语等多个语种的权威评测中位列全球第一，被 Hugging Face 等国际主流社区广泛采用。不同规模和能力的模型，为开发者根据任务、设备与成本条件进行选择提供了基础。
- **社区协作：从共同使用到共同建设。**千问的生态价值，不仅来自基础模型，也来自充满活力的社区合作。千问今天的生态价值，很大程度上并不是千问自己“做出来”的，而是社区“长出来”的。

全球 AI 模型厂商在 HuggingFace 上的衍生模型数量



来源：State of Open Models: Summer 2026 Observations, Hugging Face

在魔搭社区上的十余万个衍生模型，绝大多数都是其他开发者在千问基础上进行二次开发、微调或改造的成果。在 Hugging Face 上的 28531 个千问模型的本地部署（GGUF）转换版本中，绝大多数是由全球开发者基于 Qwen 二次开发的。这说明，千问的成功更多体现为一个由社区驱动、自我繁荣的开放生态系统，而不仅仅是一个模型产品。

- **多语言部署：为不同文化背景开发者提供可能。**普惠不仅涉及技术门槛，也涉及语言门槛。低资源语言使用者可能因训练数据不足，而难以获得与主流语言相当的服务，而多语言部署的开源模型则可能解决这一问题。

截至 2026 年 9 月，千问系列模型支持超过 200 种语言及方言，并加强了对南亚、大洋洲和非洲等地区低资源语言的覆盖。

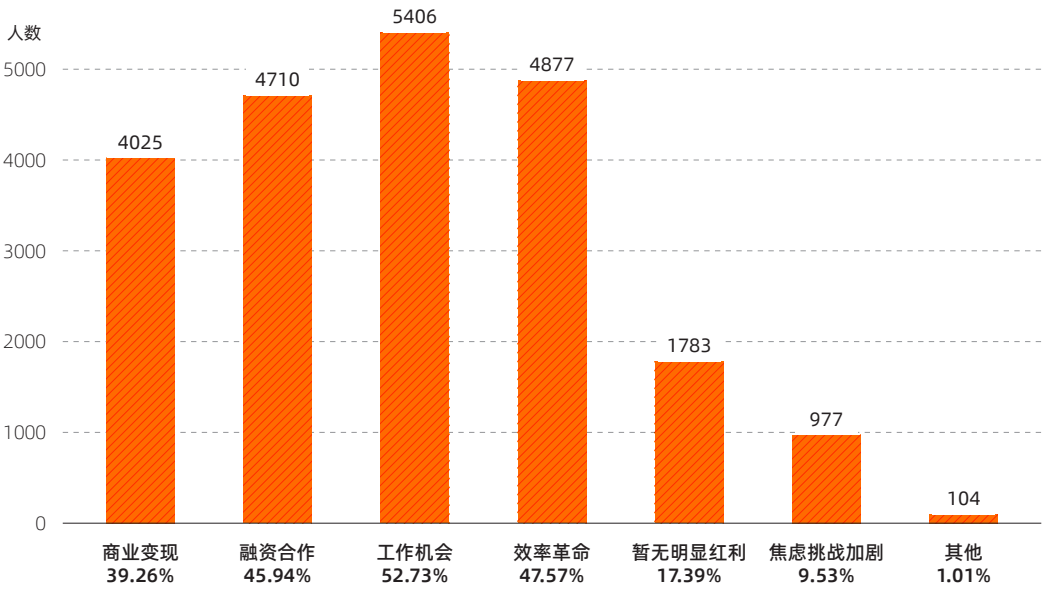
开源与安全的关系是过去一年的重要议题。诸多案例显示，安全不只取决于开源或闭源本身，而取决于模型能力、部署边界、权限控制、监控审计等一系列因素。

同时，开源也展现出有助于改善安全的特点：公开参数权重和训练方法带来的透明度，使潜在风险更容易被发现和讨论；多方参与的协同治理，有助于及时发现错误和偏差；模型规约和内生安全机制的完善，也在增强公众和监管机构对大模型系统的信任。

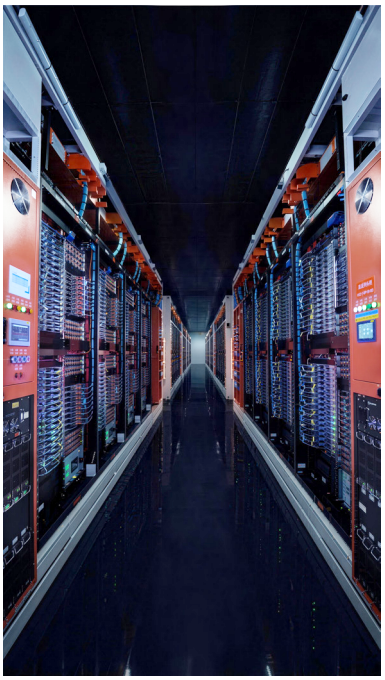
开源，开放，开拓

一项以“开源社区视角下 AI 人才发展”为主题，覆盖魔搭社区及其合作伙伴的 20 万用户的问卷调查发现，在 AI 所带来的剧烈的冲击与重构之中，开源社区正扮演着至关重要的积极角色，成为了 AI 技术扩散、应用孵化、治理共识的重要平台，也为人才培养提供了新的路径。

开源社区 /AI 对职业生涯的实际影响（多选）



来源：2026 开源人才报告，P17



绿色基础设施和算电协同

过去一年，全球 AI 基础设施建设全面加速。算力需求快速增长，对电力的需求也随之增加，能源电力已成为 AI 发展的重要瓶颈。而全球应对气候变化的压力同样在加大，即使在发电总量最大的中国，碳中和承诺的时间表也日益紧迫。

绿色低碳是阿里巴巴 AI 向善六大指导原则之一。2021 年，阿里巴巴对外公布碳中和目标。其中，阿里云作为数字化基础设施，在同阿里巴巴一起实现范围 1 和范围 2 碳中和的基础上，率先实现范围 3 的碳中和，成为绿色云。为此，我们围绕数据中心的规划、建设和运行开展全链路系统优化：算电协同规划放大能源优势，更优架构提高效率上限，运行和调度持续优化效能。

- **算电协同规划放大能源优势：**阿里云将电力供应的充沛性、稳定性、价格水平，以及清洁电力的可获取性，作为选址的重要依据，使算力布局与电力供给相协调。

  2026 财年，自建数据中心清洁电力使用比例达到 73.6%，较上一财年提高 9.6 个百分点；其中广东、江苏和四川实现 100% 清洁电力覆盖。全球租赁数据中心的清洁电力占比提升至 40.8%，同比提高 18.4 个百分点。

  同时，在能源供给相对紧缺区域，阿里云数据中心开始因地制宜地推动构建能源大基地，探索绿电直连、新型核能、海上能源岛等新模式。

- **更优架构提高效率上限：** 选址确定之后，架构设计决定了设施的能效上限。供电架构方面，阿里巴巴探索减少从电源到芯片的转换环节。基于与合作伙伴共同推出的高压直流供电架构（巴拿马电源），最新一代数据中心兼容 800V 直流供电；同时探索固态变压器（SST）等技术，为新能源高压直流直供提供技术路径。

  建设形态方面，阿里云采用 CUBE 5.0 技术架构和全模块化设计，将供电、冷却、安防、智能化、消防五大系统在工厂预集成模块，结合液冷设备支持高密度算力部署，大幅缩短了现场建设周期。整体模块化率由 30% 提升至 90%，数据中心实现 100 天敏捷交付。

- **运行和调度持续优化能效：**在数据中心运行阶段，阿里云基于数字孪生模型结合 AI 优化算法，对冷却系统运行模式及关键设备参数进行虚拟仿真和优化分析，



阿里云乌兰察布数据中心

形成最佳控制策略；并经仿真验证后下发至实际系统执行，持续提升数据中心运行能效。AI 驱动的负载优化与冷却调控已在 24 个自建数据中心落地，2026 财年自建数据中心年均 PUE（电能使用效率，即总能耗与 IT 设备能耗之比）由 1.190 下降至 1.187。

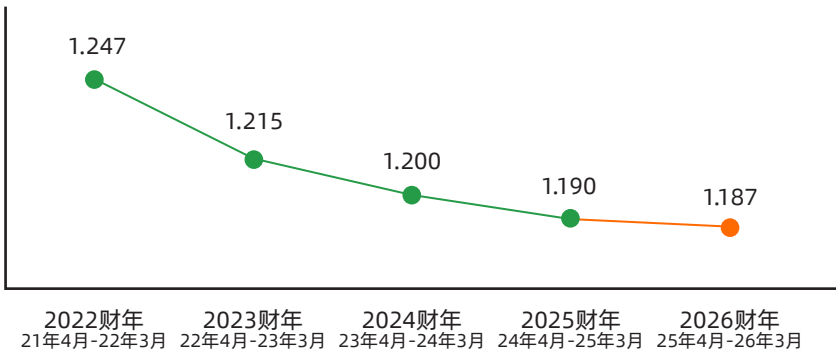
跨区域层面，考虑到 AI 负载具备时空灵活性，阿里云为此构建了跨区域算电联合调度框架，通过预测和实时动态管理机制，响应电网释放的容量与价格信号，将算力灵活调度到电力充裕的时段和地区，并参与辅助服务。数据中心因此开始演化为能源系统中的主动调节节点。

过去一年，我们还注意到数据中心冷却对水资源的消耗问题。液冷等技术减少了对蒸发散热的依赖，但随着算力密度持续提升，水资源约束需要更早纳入选址和散热方案的设计考量。

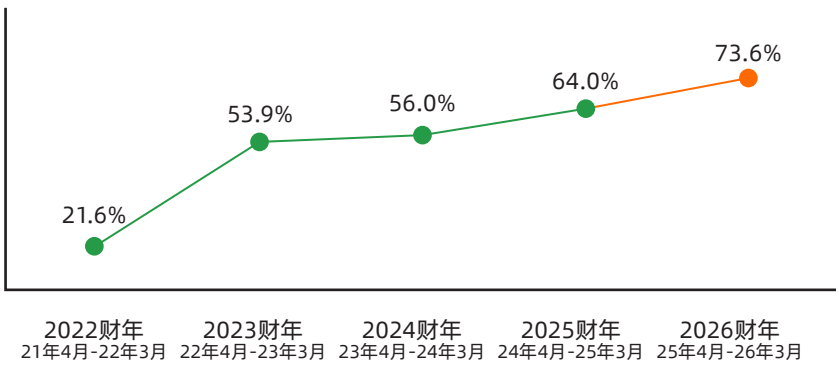
绿色 AI 基础设施的打造，仅靠个别企业不可能实现，需要全产业链的密切合作。阿里云通过发起云谷可持续论坛，与近 50 家产学研用单位共创“人工智能 + 能源”方向及算电协同前沿问题，启动联合研发课题组。同时，在海外基础设施建设中，与合作伙伴探索将全模块化的 CUBE5.0 架构和绿色能源方案结合的交付模式，根据当地能源结构和资源条件进行适配。

“过去一年，我们还注意到数据中心冷却对水资源的消耗问题。液冷等技术减少了对蒸发散热的依赖，但随着算力密度持续提升，水资源约束需要更早纳入选址和散热方案的设计考量。

过去五个财年阿里云自建数据中心的 PUE



过去五个财年阿里云自建数据中心清洁电力使用百分比




第三章
向善行动 2026
小雨正站在窗边擦玻璃，水珠顺着窗边滑下
像在玩透明的滑滑梯。

AI 为扩大向善实践的覆盖范围、降低服务成本、回应个性化需求提供了新的可能。依托全栈 AI 能力、开源与社区，以及真实场景中的应用与反馈，我们持续探索赋能组织、服务个体的方式，推动 AI 从少数先行者的探索走向更广泛的应用，让技术进步转化为可感知、可获得的实际益处。

从自身业务出发，阿里巴巴持续探索利用 AI 回应社会需求。在上一章介绍基础性工作的基础上，本章汇集自去年报告以来 AI 向善实践的新进展，围绕六个方向呈现具体应用探索及其与联合国可持续发展目标的联系。

六个方向既包括面向具体领域的探索，也包括贯穿多个领域的行动。其中，健康、教育和科研分别关注 AI 如何改善健康服务、支持学习与教学、拓展科学发现的能力，相关实践主要关联良好健康与福祉（SDG 3）、优质教育（SDG 4）以及产业、创新和基础设施（SDG 9）等目标。经济机会、安全包容的社区和绿色发展则具有跨领域特征：经济机会关注就业、创业与生产力提升；安全包容的社区关注不同群体获得技术、服务和参与机会的条件；绿色发展关注能源与资源利用、气候应对及生态保护。这三个方向分别关联多个可持续发展目标，也与健康、教育和科研领域的实践相互交织。

与去年相比，今年报告收录的案例覆盖了更广泛的可持续发展议题，也呈现出更加多元的参与主体和服务对象。从这些案例中，我们看到一些新的探索：行业经验通过 AI 转化为更易调用和复用的能力，科研工作借助 AI 探索新的知识发现路径，AI 应用也逐步从单一工具走向组织流程，融入更多生产与生活场景。

本章不仅呈现阶段性成果，也记录仍在推进的探索：一些案例已取得初步成效，一些仍在验证可行性，还有一些正根据实践反馈调整方向。我们希望分享其中的经验与教训、机会与挑战，并与更多伙伴共同探索，让 AI 向善从技术的可能走向持续的行动。



AI 赋能健康

AI for Health

达摩院平扫 CT+AI
阿里健康“氢离子”
追星星的 AI

全球医疗体系正同时面临需求快速增长与供给持续紧张的双重压力。在中国，从疾病筛查、临床诊断，到药物研发和特殊人群干预，医疗健康服务链条的不同环节普遍存在专业资源不足、服务效率有限、优质资源分布不均等问题。

AI 的加入，有望提高专业医疗服务效率，并扩大整体医疗供给。预测模型、决策支持工具和生成式工具，已经开始进入临床医疗与日常照护。但技术上的可行，并不等于实践上的可用。例如，医疗服务高度本地化、碎片化和强监管的特点，决定了 AI 赋能健康必须从设计、训练、部署到评估都要因地制宜。患者、医生、医院、保险、药企、监管等多方主体，需要在临床安全、组织流程和商业激励之间建立新的平衡。

达摩院平扫 CT+AI

医疗基础设施的低成本、高价值进化路径

平扫 CT 是临床最普及、最标准化的设备之一，每年产生海量医学影像，却长期没有充分释放其价值。过去，一张 CT 通常服务于一次特定疾病诊断，医生只关注单一病灶，同一影像中大量疾病信息未被系统利用。同时，即使针对同一病变，人眼阅片在分辨能力上的天然局限，也限制了影像价值的进一步挖掘。

AI 让平扫 CT 具备了持续进化的能力。阿里巴巴达摩院的“平扫 CT+AI”解决方案，在不改变设备和诊疗流程的前提下，不断扩展同一张影像能够回答的问题，使一次常规检查逐步演进为覆盖筛查、风险评估和辅助决策的综合性医疗健康入口，推动 AI 医疗从单病种筛查走向“一扫多筛”。

2026 年，“平扫 CT+AI”的主要进展，是多癌种筛查能力的持续提升。同时围绕同一张平扫 CT，达摩院的能力已逐步扩展至急症预警、慢病评估等，形成持续演进的医疗 AI 能力体系。

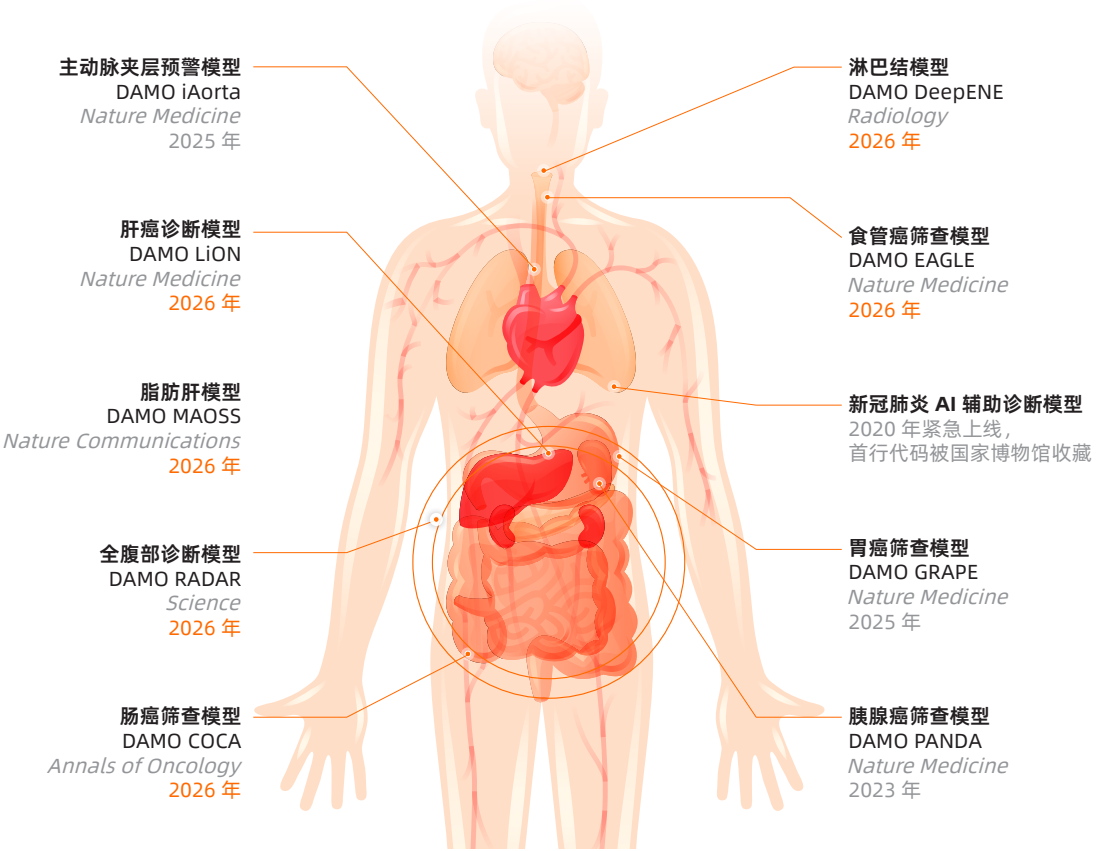
达摩院早在 2023 年就发布了胰腺癌筛查模型，在国际上首次验证了平扫 CT 胰腺癌早筛的可行性，《自然·医学》称之为“基于医疗影像 AI 的癌症筛查即将进入黄金时代”。过去一年，胃癌、结直肠癌、肝癌与食管癌筛查模型相继发布，进一步证明传统上被认为难以通过平扫 CT 识别的脏器，同样可能蕴含丰富而可识别的疾病信号。

与去年相比，在癌症领域，达摩院还在探索乳腺癌、肾癌等更多癌种；在急症领域，肺栓塞等危急重症识别能力也在持续扩展；在慢病领域，骨质疏松、脂肪肝等模型已开展应用，瓣膜病等也启动探索；与此同时，DeepHTD 模型已尝试预

阿里巴巴也在探索通过“AI+ 语音辅助”筛查疾病

2026 年，在第二届“小有可为”AI 向善创新挑战赛上，阿尔茨海默病 AI 早筛项目“勿忘我”获得普惠养老赛道一等奖。该应用借助 Qwen-Omni 全模态能力，通过语速、停顿比率、语义连贯性、应答反应等语音及声学特征，能在一通电话中完成认知风险的初步评估。

达摩院 AI 识别 CT 细微病灶，突破全球重大医学难题



测肺癌患者接受新辅助治疗后的病理缓解程度，为临床提供辅助参考。

不同模型覆盖不同疾病，却建立在同一条技术路线之上。随着可识别疾病不断增加，同一张 CT 新增一种筛查能力的边际成本持续下降，既有医疗基础设施的价值不断提升。更重要的是，这一能力增长无需增加患者检查次数和辐射暴露，也无需改变医生既有诊疗流程。

9 月，达摩院开源了通用医疗影像 AI 模型 DAMO RADAR。它面向腹部增强 CT 诊断，覆盖 18 个器官，可一次性识别超过 146 种病症。它区分患者与正常人的水平（即 AUC 指标），首次达到了影像科专家水平。在它的提示下，影像科医生的防漏诊能力提高了 10%，用时则减少了 30% 以上。这一成果登上了《科学》杂志。

更通用的筛查技术，可以将多项检查的高昂费用，降低到一次平扫 CT 加 AI 解决方案的少量费用。同时，骨密度或脂肪肝等机会性筛查也成为可能，原本不会主动进入筛查流程的人群获得了早期识别和干预机会。

中国国家癌症中心年初刊发在《自然》子刊上的文章指出，中国人口约占全球 17%，却承担了约 25% 的全球癌症负担，癌症患者五年相对生存率仍低于美国等发达国家。早筛是应对癌症的关键，不仅可以挽救更多生命，也有助于节约医疗成本。

然而，AI 在医疗体系的规模化部署，还取决于能否建立一条从数据获取、模型训练到临床验证的持续演进路径。新冠疫情期间，全球曾涌现数百种 CT 影像 AI 工具，但疫情结束后，大多数产品并未实现规模化应用。英国图灵研究所复盘发现，训练数据质量不足、缺乏真实世界验证以及未融入临

床工作流，是其共同原因。

中国拥有丰富的医学影像资源和庞大的患者群体，大型医疗机构持续积累真实病例；同时，相比高度市场化的医疗体系，国内跨医院协同开展临床研究、建设高质量数据集和推进真实世界验证，也具备更低的组织成本。这成为在中国不断拓展“平扫 CT+AI”能力边界的重要基础。

真正有效的医疗 AI，并不是要求医生适应 AI，而是让 AI 适应医生既有工作流。急性主动脉综合征便是典型例子。该疾病进展迅速，漏诊和误诊风险较高，而超过一半患者初次就诊时接受的是平扫 CT。

针对这一临床痛点，DAMO iAorta 模型连接医院影像信息系统，可在几秒内实时判读平扫 CT 图像，一旦发现高危病例，就推送至医生界面提醒优先查看。模型同时生成可疑区域定位热力图，让医生能够直观理解 AI 判断依据；它也成为医生持续学习和积累经验的新工具，加快年轻医生成长，促进医疗经验在不同层级医疗机构之间传播。

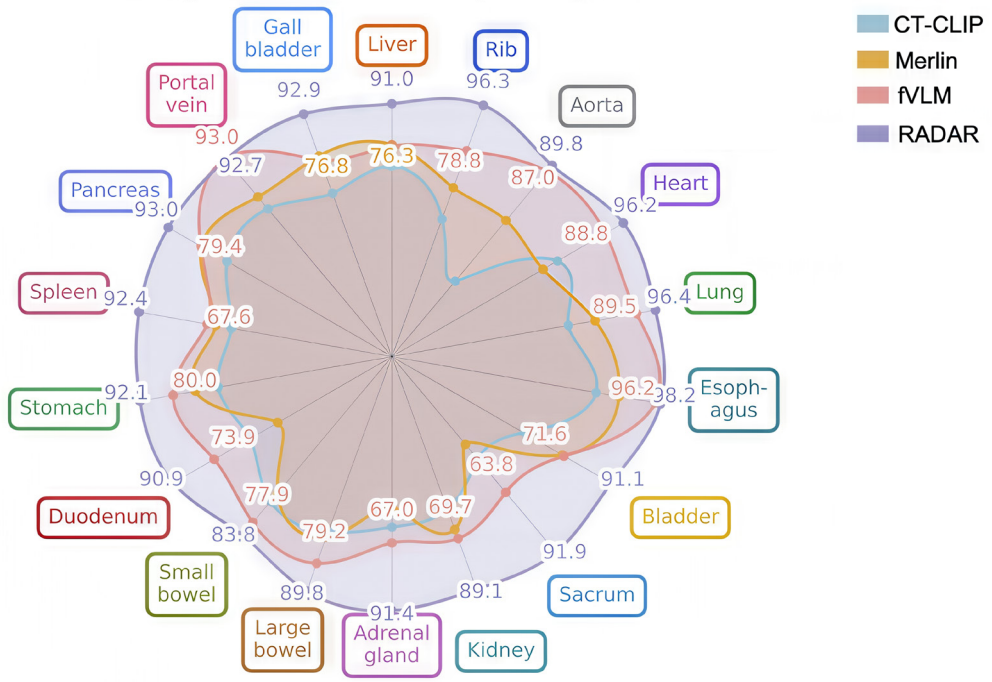
这一价值在基层医疗体系中尤为明显。围绕该模型，浙江绍兴某医院试点构建了 15 分钟抢救圈，分院上传 CT 片，由中心医院 AI 实时预警。试点数据显示，平均确诊时间从 4.3 小时缩短至 1.7 小时，低年资医生诊断敏感性提升约 40%。医疗供给不再完全受限于专家数量，而开始通过算法能力实现更大范围、更低成本的扩展。

AI 赋能健康必须经过多轮严格的临床循证验证。医疗 AI 需要持续积累前瞻性临床研究和真实世界数据，证明其能够提高早期发现率、改善诊疗决策，并最终转化为患者获益。这一过程并非一次性完成，而是在真实应用中不断积累证据、完善模型。

越来越多医疗机构正在共同参与医疗 AI 循证证据的持续积累。2026 年，达摩院已和 GE 医疗、东软医疗等国内外头部 CT 机厂商签约，并和多家医疗机构达成合作；国内首批 10 家医院已部署主动脉急诊 AI 模型 DAMO iAorta；多癌筛查模型累计为全球 11 个国家 / 地区提供超过 2000 万人次 AI 服务。我们看到，从算法到证据再到制度，可能是支撑 AI 医疗创新的一条新路径。

通用医疗影像 AI 模型首次达到专家水平

在 146 种医学体征与疾病诊断任务中，DAMO RADAR 识别患者的表现全面领先



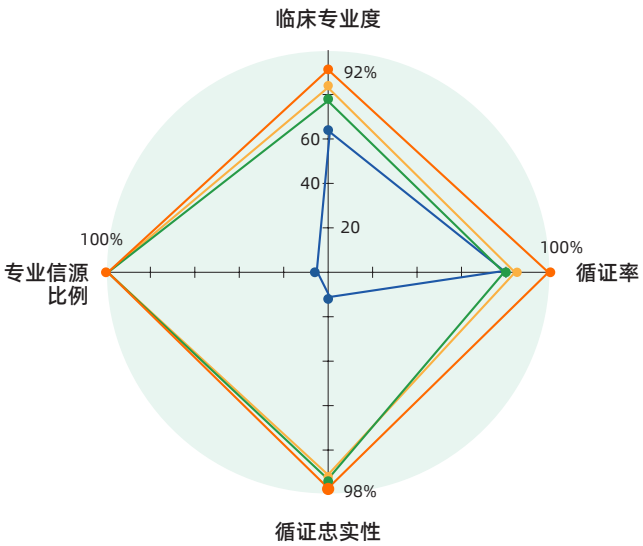
阿里健康“氢离子”

以循证赢得医生信任，以 AI 推动诊疗进步

医疗决策的本质是在有限时间内对无限医学证据进行组织与判断。中国拥有 500 万名医生，每年完成约 100 亿人次诊疗。每一次诊疗都是一次对医学知识的个体化重组，而这些决策质量构成了医疗体系运行的基础。

2026 年 5 月，阿里健康推出面向医生群体的 AI 原生应用“氢离子”，探索的正是一套循证医学驱动的决策支持体系。其目标不是替代医生判断，而是让已有医学知识更可靠、更高效地进入临床决策，帮助医生围绕具体病例完成证据检索、鉴别分析和决策验证。

氢离子在循证医学相关基准上领先同行





说明：临床专业性指关键观点、论据与结论的准确率；循证率指答案段落级引用证据的比例；循证忠实性指观点、论据、结论与引用证据内容的一致性；专业信源比例指引用证据中专业信源的比例。

这一探索回应的是现代医疗体系面临的一个核心矛盾：医学知识持续增长，疾病谱系不断细化，长尾病例不断出现，而医生利用知识的能力仍受到时间、经验和信息获取方式的限制。截至 2026 年 9 月，全球公开医学文献已超过 4000 万篇，且更新速度仍在加快。长期以来，医学知识主要依赖论文、指南和学术会议传播，医生需要在有限时间内完成检索、筛选、理解和应用。资源有限的地区面临的压力更大。在此背景下，诊断错误和医疗决策失误仍是现代医疗体系的重要挑战。

与此同时，通用大模型虽然具备较强的语言理解和生成能力，却难以满足医学场景对准确性、可靠性和可溯源性的要求，可能生成看似合理但缺乏证据支持、与事实不符甚至虚构的内容，即所谓的“幻觉”；还可能受训练数据时效性、专业深度和推理稳定性不足的限制。因此，通用大模型不能直接作为临床决策依据，也说明我们需要更可靠、可验证、能结合专业知识的医学智能系统。

正因为此，对于医疗体系而言，可信赖的 AI 比更聪明的 AI 更重要。医疗 AI 只有建立证据、形成信任，才能从实验室进入诊室，从技术能力转化为医疗能力。因此，它必须具备可追溯的证据来源、可解释的推理过程，以及持续验证的临床反馈。

截至 2026 年 9 月 30 日，氢离子已先后与四大国际医学顶刊中的 BMJ 集团、NEJM 集团、JAMA Network，以及国内的中华医学会杂志社、人民卫生出版社、中国抗癌协会等权威学术机构达成内容合作，并联合顶尖行业专家成立医学 AI 专家委员会，共建医学 AI 数据集。

专业信源只是基础，更重要的是建立循证驱动的推理机制。氢离子将 AI 对医学证据的使用约束在循证医学框架内，为医生同步呈现证据来源、证据等级、适用范围及局限性。目前，氢离子在医学场景中实现 92% 准确率、100% 循证率、98% 循证忠实性，并通过医生的真实使用反馈，持续优化临床体验。约 30% 的新增医生用户来自同行推荐，也印证了医疗 AI 正经历从技术认可走向临床信任的关键转变。

随着规模化应用，氢离子正推动医学 AI 从单点技术验证迈向医学循证场景的全域应用。过去，前沿医学进展主要依赖专家、期刊和学术会议进入临床，存在滞后和覆盖不均等局限。未来，AI 有机会让经过验证的医学证据以更低成本更高频次嵌入日常诊疗流程，增强医生的持续学习与临床决策能力，促进优质医疗资源可及性与健康公平。

追星星的 AI

特殊人群干预，重新连接自我与世界

孤独症谱系障碍是一种神经发育障碍。在中国，0-14 岁孤独症儿童规模已超过 200 万，每年新增约 20 万，亟待干预治疗。但其干预方式高度依赖儿童自身的兴趣偏好、认知水平和行为特征，因而面临开发成本高、专业门槛严、个性化要求细的三重约束。

在阿里巴巴公益支持下，2024 年阿里巴巴志愿者联合特殊教育专家推出的“追星星的 AI”绘本智能体，探索用生成式 AI，缓解这个残酷的“不可能三角”。家长或教师只需输入简单指令，绘本智能体即可调用千问系列大模型完成相应故事、图片、语音等生成任务，并在智能体编排下，生成可动态调整的有声绘本。

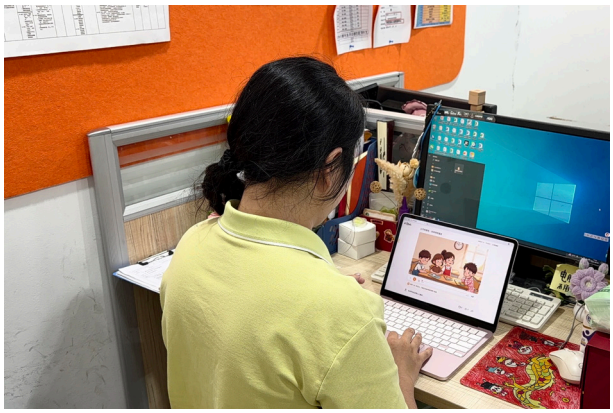
2026 年，“追星星的 AI”完成模型层升级。长文本理解与生成模型 Qwen-long，可完整摄入专业特教指导文档、经典绘本结构模板、情绪发展心理学资料等上下文；并在垂类语义理解上做了针对性优化，使生成内容兼顾心理学规范与儿童接受度。多模态画面引擎 Qwen-image 确保角色在跨页间的一致性。两款模型协同实现“文图语义对齐”，让每一页文字和画面服务干预目标。

这是典型的 AI 原生应用。模型能力的进步不仅提升了效率与质量，更创造了过去不存在的用户体验。过去一年，语音复刻的样本门槛从分钟级降至 4 秒钟。这让原本难以持续完成表达的孤独症儿童，也能参与内容生成，从倾听他人转为讲述自己。AI 降低了特殊人群平等参与日常生活的门槛，重新连接自我与世界。

截至 2026 年 6 月，“追星星的 AI”绘本智能体累计提供的绘本生成和阅读服务已超过 40 万次，是去年同期 2 倍，已覆盖孤独症儿童及特教老师超 2 万人。

AI 原生的产品技术架构

百炼 Agent 平台	阿里云 & 魔搭社区
故事生成 Agent	Qwen-doc-turbo
图像生成 Agent	Qwen-image
语音生成服务	CosyVoice2
Comfy UI 工作流	LoRA 模型训练



特教老师使用“追星星的 AI”辅助备课。

展望：

我们正在进入医疗健康领域 AI 规模化应用的新阶段。正如强监管的自动驾驶领域历经 20 余年的严谨迭代，才逐步展现出超越人类驾驶员的性能，医疗健康 AI 同样需要在持续验证中建立信任。未来的挑战，除了在已有充分证据支持的场景中加速应用，同时还要通过临床试验填补关键证据空白，才能推动医疗 AI 从局部验证走向大规模普及。

高质量数据仍然是医疗 AI 规模化发展的关键基础设施。未来，医疗数据需要从分散资源转变为可安全共享、可持续利用的基础设施。通过多模态数据融合、隐私保护计算和持续反馈机制，医疗系统有望形成数据、模型与临床实践相互促进的演进闭环。

AI 赋能教育
AI for Education

少年云助学计划
云谷学校
千问高考志愿填报智能体
高校 AI 公益支持计划
乐云 RynnBot 平台
阿里云科普教育基地

经过多年发展，中国已建成世界上规模最大且有质量的教育体系。在取得成绩的同时，教育领域的发展挑战依然存在。例如，优质资源仍受到地域、师资与基础设施等因素制约，教育公平依然存在提升空间。又如，传统教育模式和评价体系过于强调知识获取与标准结果，挤压了个性化学习、批判性思维与综合能力培养的空间。

随着技术快速迭代，教育体系还需要回答一个新问题：如何让不同年龄、不同地区的人，都获得理解和使用新技术的能力。

AI 正在改变教育资源与教育能力的供给方式。从基础教育的个性化学习，到升学场景的智能决策，再到高校的教学科研，AI 让教育服务能够兼顾个性化与规模化。但随着 AI 能力本身成为一种新的基础能力，并改变未来社会的人才需求结构，教育与 AI 的结合，需要让更多人获得驾驭 AI 的能力。如何帮助教师、学生和学校主动重构学习过程，正在成为 AI 赋能教育的新课题。

少年云助学计划

教育公平从资源可及走向能力可及

AI 具有强大的知识生成能力，能用于各种教育水平、不同发展阶段和文化背景的学习者。但这种能力是让教育鸿沟进一步扩大，还是真正实现教育平权，很大程度上取决于教育体系培养学生用好 AI 的能力。教师正在从知识传递者转向学习过程的设计者与引导者，学生则从被动接受答案转向主动提问和独立判断。

阿里巴巴“少年云助学计划”启动至今已有 5 年，鉴于教育公平正从资源可及性走向能力可及性，项目也在不断迭代发展。截至 2026 年 3 月，少年云已覆盖全国 28 个省份，建设了 403 间 AI 云教室，支持师生拥有 38 万个个人云电脑账号。

项目学校中有 5 所入选教育部中小学人工智能教育基地，并获得联合国教科文组织等相关机构颁发的“全球世界遗产教育创新案例奖”、“全球智慧教育创新奖（实践创新）”等。

AI 进入教育的挑战之一，在于推动教师角色、教学流程和评价体系同步转变。当前，教师更多作为工具使用者参与，缺少持续参与课程设计和实践创新的机制。因此，少年云将着力点转向教师能力建设。

2026 年，在新增 100 间云教室的同时，项目发起了“万名 AI 先锋教师计划”，为全国教师提供 AI 教学实践支持，并建立跨区域教研共同体；进一步引入国内头部学校，以共建共享模式，在不同地区 AI 教育中探索起点公平机制。

在课堂实践中，AI 正在改变学习的开展方式。少年云的 AI 智能体能够分析学生的学习轨迹、能力分布和高频错误，使原本不可见的学习过程变得可记录可分析，为因材施教提供数据和技术支持。AI 智能体也为学生提供了低门槛的互动环境，让学生更愿意提出问题并探索未知。当答案越来越容易获得，提出问题本身成为新的学习能力。

在课程研发层面，少年云与北京大学教育学院等合作，尝试研究学生如何理解与评估 AI，以及与 AI 互动过程中可能出现的认知偏差、过度依赖等问题，并据此优化课程设计。

云谷学校探索 AI 教育深水区

AI 如何促进学习而非替代学习

随着 AI 深度融入教育，如何判定 AI 技术和学习过程之间的边界和分工，成为教育者必须把握的关键：AI 介入后，真正的学习有没有发生，人本身的能力有没有变强，学生是否形成理解他人、与真实世界相处的能力。

由阿里巴巴合伙人投资创办的云谷学校，在过去一年深度探索 AI 教育时，逐步正视并尝试应对 AI 进入基础教育的四大新挑战：

- 从关注 AI 提高多少产出，转向关注学生真正学会了什么，警惕学生产生“能力幻觉”；
- 从单纯追求效率，转向守住阅读、思考、试错等必要的“低效”过程，避免“认知外包”；
- 从评价最终结果，转向评价思考过程，尤其关注学生在离开 AI 后，能力能否迁移；
- 从讨论 AI 如何增强知识认知，转向兼顾青少年的社会化，尤其是帮助孩子学会理解真实的人，以及处理与社会关系的摩擦与修复。

基于这些思考，云谷学校也在从零散的工具实践，逐步走向在 AI 时代对教师能力、学生能力、课程体系和学校定位的系统重构。

首先，建立教师的技术判断力。在这个过程中，“老师要快”，即更快理解、试用和识别技术风险，判断 AI 有没有真正促进学习；而“学生要慢”，即孩子的成长节奏不能被模型迭代速度所裹挟。

其次，培养学生的 AI 交互主动性。在使用 AI 前，学生应先思考、拆解任务，决定哪些交给 AI、哪些由自己完成，并持续判断结果、发现问题和调整策略。让 AI 从“答案机器”变成“思考伙伴”。

再次，重新设定学习任务与评价。积极使用 AI 获取资料、反馈和学习机会，但保护阅读、思考、写作、动手，以及失败和修改的学习过程。评价基准也相应调整，从只看最终产出，转向收集必要的过程证据。

最后，定位学校为“真实关系的练习场”。技术越虚拟，教育越要创造真实连接；AI 越懂孩子，孩子越需要学习去懂一个真实的人。

2026 年，云谷学校中学部设立浙江省首个人工智能创新应用特色班，尝试用高中三年时间构建 AI 人才培养方案。这套方案更强调从一个 15 岁孩子三年后的成长目标倒推课程设置，包括应该具备怎样的科学基础、工程能力、问题意识、AI 能力和创造经验。



云谷教师先后出版了《从工具到伙伴：教师 AI 实践案例》、《用 AI 解锁中小学科技教育》

千问高考志愿填报智能体

弥合信息资源和认知能力鸿沟

填报志愿是教育公平的“二次分配”，也是学生和家长最焦虑的时刻之一。每年，超 1000 万考生需要在短时间内，从近 3000 所高校、2000 多个专业中完成高度复杂的组合决策。家庭信息资源和认知能力差异带来的影响，在基础教育阶段可能已被部分弥合，却可能在志愿填报这一刻重新放大。

2026 年，千问上线国内首个全周期高考志愿填报智能体，为全国考生免费提供志愿填报和咨询服务。与 2025 年的“夸克高考大模型”相比，这款智能体更具自主性与长期记忆能力。

用户与智能体的交互覆盖考后规划、查分准备、正式填报等关键阶段，持续陪伴考生完成志愿决策。同类服务市场收费往往高达数千元，而全国仅 5% 考生聘请了付费专家指导服务。

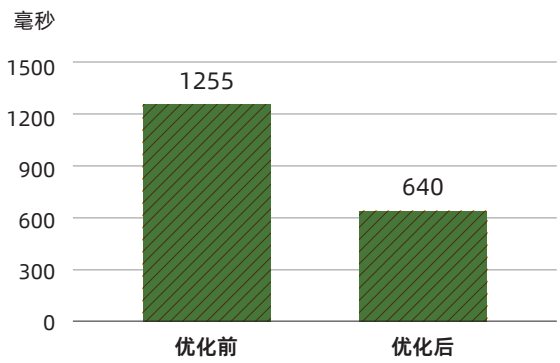
在去年的实践中，团队发现 35% 用户只填了省份、选科、分数、位次四个基本要素，对新兴产业与未来职业缺乏足够理解，也无法提出更深层的追问。实际上，许多考生真正缺失的，并不只是信息本身，更在于构建决策框架和提出关键问题的能力。

今年，千问高考志愿填报智能体开始更深地介入志愿填报的决策过程。它强化了动态调整与主动建议功能，通过多轮对话，帮助考生不断完善自我画像、明确真实需求，并逐步缩小选择空间。所有推荐结果均保留可追溯依据，由考生做出最终决策。

此外，千问工程团队还针对老旧机型与网络条件较弱的乡村地区进行专项优化，并联合“暖芒公益”计划，进一步缩小县域及乡村地区考生在校园之外面临的资源差距。

智能体辅助填报志愿，相当于引导年轻人第一次系统认识自己、理解专业与职业。2026 年高考季，用户在千问上共领取超过 2300 万份免费 AI 志愿报告，是去年 1300 万份的近 2 倍。

低端机性能优化，
打开千问高考智能体速度提升 1 倍



高校 AI 公益支持计划

探索面向未来的人机协作能力

AI 正在重塑高等教育的教学与科研，也在改变 AI 时代市场对高校人才培养的期待。高校可能成为年轻人最早接触并适应“白领技术革命”的场景。然而，前沿 AI 工具的获取仍存在成本、语言和网络等门槛，不同学生之间的 AI 使用能力可能进一步拉大。

2026 年，阿里云与阿里巴巴公益发起“高校 AI 公益支持计划”，为高校大学生提供两款免费的 AI 工具：千问办公和 Qoder。这为学生提供了一套面向未来知识工作的 AI 生产工具。

该项目第一年计划覆盖近 10 所双一流高校的超过 20 万在校学生（本科、硕士、博士），激活坐席数 10 万个。9 月 19 日，项目首站落地中国科学技术大学，并现场聘任了“千问办公校园大使”“Qoder 探索大使”，他们将在项目实施过程中为更多同学们提供支持。

9 月 24 日，计划落地上海交通大学，阿里面向所有在册师生提供一年期免费授权及相应算力额度，并配套产品培训、技术支持和校园活动。截至本报告发布，更多其他高校正在陆续开通中。

从开放工具到支持实践，再到分享方法，高校 AI 公益支持计划希望帮助学生不仅成为 AI 工具的使用者，更成为创造者。其长期目标，是希望帮助大学生不只利用 AI 工具完成更多任务，而是在保有独立思考与判断基础上，逐步形成面向未来知识工作的人机协作能力。



千问办公校园大使、Qoder 探索大使聘任仪式

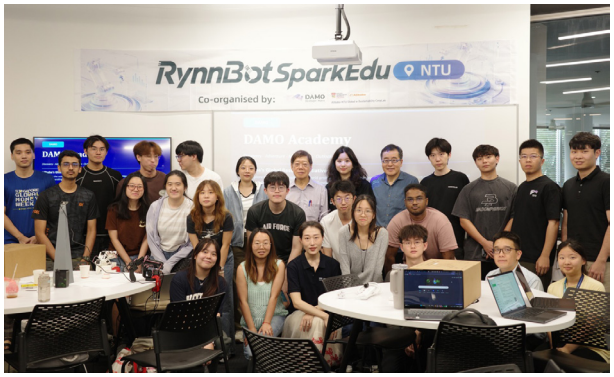
乐云 RynnBot 平台

培养 AI 前沿交叉领域复合型人才

具身智能是 AI 与工程深度融合的前沿方向。具身智能开发不仅涉及机器人本体，还需要完成数据采集、模型训练、仿真验证和真机部署等多个环节，技术门槛和硬件投入较高，高校学生往往难以获得贯穿完整开发流程的实践机会。

为降低这一门槛，阿里达摩院于 2025 年推出乐云具身智能平台（RynnBot），并开源机器人上下文协议 RynnRCP；2026 年 2 月，进一步开源具身智能基础模型 RynnBrain。借助端云协同，开发者无需在端侧配置 GPU，即可运行主流具身模型。

截至 2026 年 7 月，面向教育场景的“乐云具身 SparkEdu”训练营已与国内外 18 所高校开展合作，伙伴关系拓展至中国香港、新加坡、巴基斯坦等地，累计聚集 1.7 万名具身智能开发者，并孵化出 100+ 个创新作品。学生的实践贯穿遥操数据采集、模型训练、真机部署全流程，为高校培养适应未来产业需求的复合型人才提供了实践路径。



乐云具身 SparkEdu 走入校园

阿里云科普教育基地

用 AI 科普 AI

AI 科普的价值，不仅在于降低使用 AI 的门槛，更重要的是让不同年龄、职业和地域的人群，都获得理解和学会新技术的机会。

阿里云科普教育基地是一个上千平方米的沉浸式 AI 展厅，员工志愿者和合作高校志愿者充当科普生力军，向社会公众提供展厅参观、科普讲座和 AI 工作坊等公益活动，让公众通过看得见、用得上的体验方式理解 AI。

针对不同人群，志愿者们设计了不同的科普方式：

- 对城市中小学生及亲子家庭，通过公益研学激发青少年的兴趣与创造力；
- 对偏远地区青少年，通过定向邀请和专场活动，让其亲身体验 AI；
- 对视障等失能群体，通过探索 AI 无障碍交互体验，探索技术助残的可能性；
- 对极客创客与高校开发者，以 AI 工作坊和开放实验形式深化技术实践。

基地在科普中全面践行“用 AI 科普 AI”，用大模型辅助生成讲解词、制作课件和设计互动环节，提升科普效率和体验。截至 2026 年 9 月，基地科普志愿者已超过 80 人，累计接待 10 万余名参观者，举办公益科普活动上百场，来访者年龄跨度从 6 岁到 78 岁。基地被评定为浙江省科普教育基地、浙江省新质生产力科普体验馆（点）。

展望：

AI 赋能教育尚未建成系统的评估框架，这一框架要回答的核心问题是，学生应该形成什么样的能力。评估体系应该验证 AI 是否放大了学习能力，而非仅关注任务完成效率，或考分的提升，这会偏离教育培养学生独立思考和持续学习的根本目标。

评估框架解决的是“评什么”的问题，而落地最终取决于教师。目前技术更新速度远快于传统教师培养进度，教师正成为 AI 教育转型中压力最大的群体之一。未来，在专业能力持续精进的基础上，教师不仅需要掌握 AI 工具，更需要组织学习过程、引导人机协作。AI 赋能教育需要建立可持续、可推广的教师培养机制。

AI 赋能科研
AI for Science

Luca 系列模型

灵枢细胞与归元模型

ElementsClaw

科学问题日益复杂，研究对象从单一学科走向跨学科、多尺度和多模态，科研人员需要处理的知识与数据指数级增长。从全球来看，尽管科研投入持续增加，但创新效率在部分领域的增长却有所放缓。中国正在进入更加前沿的科学发现、更加底层的技术研发新阶段，亟需探索新的研发范式。

AI 正在重构科学发现的方式，降低了跨学科研究门槛，使研究者能够更高效地整合不同领域的知识与数据，加速从假设到验证的研发进程。目前，AI 已逐步进入生命科学、材料科学，新能源和气候变化等领域。研究机构在训练不同科学领域的专有模型的同时，也在构建面向科学发现的智能体，让实验反馈参与下一轮研究，甚至推动模型自身迭代。借助 AI 推动科学发现的重大突破，是 AI 向善最值得期待的领域之一。

Luca 系列模型

汇入科学发现工作流

AI 赋能科学正在进入一个新阶段，通用大模型、科学智能体、自动化代码执行与自动实验验证等能力，正在与科学家共同塑造新的科学发现范式。在这一新范式里，AlphaFold 等专用 AI 模型仍然不可替代，它们将成为通用科学智能体可以调用的专业能力模块。

2026 年，阿里巴巴 Qwen-3.8-Max 已经实现从零构建论文中的完整训练与评测流程，复现了论文的六项主要结论，还像人类研究员一样独立提出多种改进方向并加以测试，最终找到超越论文原方法的新方案。这一能力依托于模型对科研目标的理解、跨学科知识的语义推理，以及智能体架构对科研任务的拆解与工具协调。

阿里巴巴的生物大模型 LucaOne，是这套新的科学发现体系下可供调用的专业工具。其训练数据覆盖 16.9 万个物种，涉及核酸序列约 12 亿条、蛋白序列约 6 亿条，是围绕生命中心法则、统一建模的生物语言模型。截至今年 6 月，LucaOne 在 HuggingFace 上的月均下载量超过 5000 次，在同类模型中表现活跃。

基于 LucaOne，Luca 系列专有领域模型家族

LucaProt
RNA 病毒鉴定模型
《细胞》
2024 年

LucaPCycle
磷循环蛋白家族发现模型
《自然·通讯》
2025 年

LucaVirus
病毒领域模型
《国家科学评论》
2026 年

LucaCell
单细胞模型
GitHub 开源
2026 年

序列是生命通用的共同语言。基于 LucaOne 的通用表征能力，团队进一步为“生命的语言”构建了 Luca 系列专有领域模型。这些专用模型开始能与 Qwen 等通用大模型协同编排、组合调用，汇入科学发现工作流。

2026 年 9 月，Luca 家族最新成员 LucaCell 开源。它同样构建在 LucaOne 序列表征能力之上，是一款跨物种、跨模态的单细胞基础模型。它在包含 8500 万个人类和小鼠单细胞转录组数据上预训练，涵盖了 62 种组织来源和 18 种疾病状态。在物种细胞鉴定、基因表达重建、病毒载量预测等任务上，LucaCell 展现了出色的泛化能力，有望为未来微生物与宿主相互作用及其细胞层面的致病机制研究提供基础。

在此基础上，阿里继续围绕前沿科研需求，推动专业工具扩展与迭代。新一代的 LucaGen 具备统一的生物序列理解与生成能力，将由满足科学发现，迈向满足科学创造；而 LucaProtGen 将拥有用自然语言设计生物蛋白的科学能力。

灵枢细胞与归元模型

生命科学从静态分析走向动态模拟

AI 正在改变药物分子的发现与设计。但新药研发真正的瓶颈，往往不在分子设计本身，而在于对疾病的理解。这要求建立患者临床表现与药物在细胞层面的作用机制之间的联系。

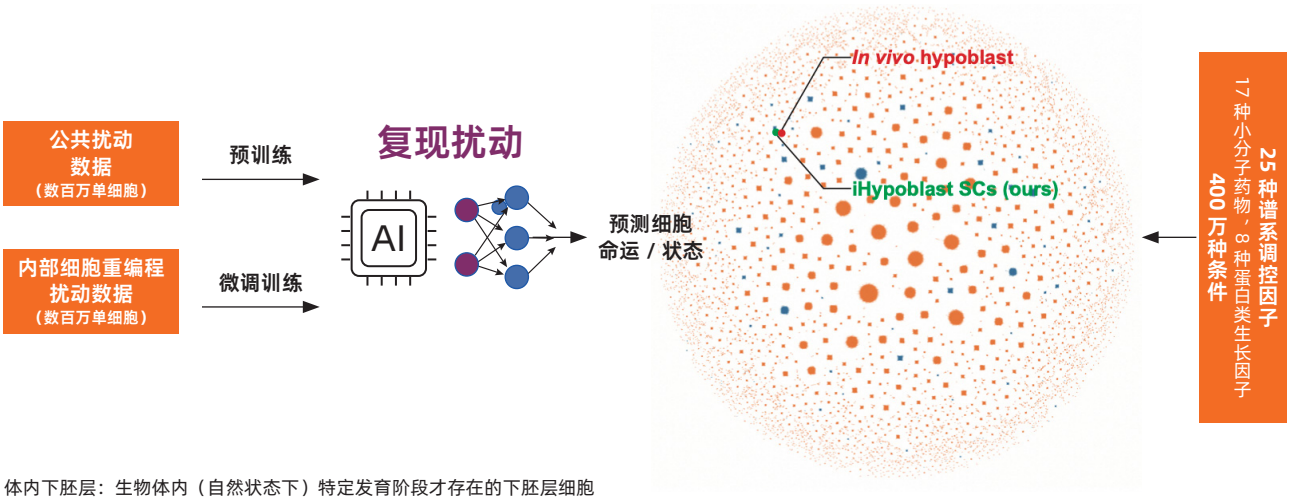
现有 AI 模型大多聚焦于分子层面的静态表征或局部相互作用，难以在更大时空尺度上模拟细胞在遗传或环境扰动下的整体响应；而在新分子或新靶点发现后，临床前研究仍高度依赖动物实验，其周期长、成本高、通量低，已成为制约新药管线推进的“卡脖子”环节。

为了突破这些瓶颈，2025 年以来，全球科研机构加速探索虚拟细胞（Virtual Cell）。阿里巴巴达摩院同样围绕这一方向，联合西湖大学等高等院校及研究机构，先后发布了生成式细胞世界模型灵枢细胞（Lingshu-Cell）与干细胞重编程预测模型归元（ReproPerturb），逐步构建起从细胞状态模拟到细胞命运预测的能力体系。

灵枢细胞能够生成不同遗传扰动和药物刺激下的细胞转录组状态，实现从静态细胞图谱向动态细胞世界模型的演进；归元则进一步融合小分子、生长因子等多模态信息，预测不同扰动组合对干细胞重编程方向的影响，并通过可解释机制揭示潜在生物学通路，为实验设计提供依据。

随着虚拟细胞不断逼近真实细胞行为，未来大量药物筛选、疾病机制研究和再生医学实验有望首先在数字世界完成，再进入实验室验证。这不仅能够缩短药物研发周期、降低研发成本，也将推动生命科学逐步形成“AI 模拟–实验验证–模型迭代”的新型研发范式。

团队通过归元模型预测的最优方案，成功获得高质量的下胚层样干细胞



体内下胚层：生物体内（自然状态下）特定发育阶段才存在的下胚层细胞
诱导下胚层干细胞（本研究的成果）：在分子特征上与天然下胚层细胞高度相似，稳定表达关键多能性因子；在体外持续传代 50 代后，它们仍保持干细胞功能特性。

ElementsClaw 超导材料发现智能体

参与组织科学发现

AI 赋能科研正从预测范式迈向生成范式，但两者都面临共同的决策瓶颈：数百万个候选材料，究竟哪个才真正值得进入实验室合成？这一决策需要协调原子尺度的数值计算（结构、稳定性）与高层语义推理（新颖性、合成可行性）等多个维度。

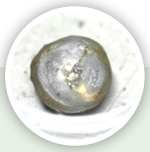
智能体提供了一条关键路径，把材料发现从孤立过程提升为集成框架。2026 年 7 月，阿里巴巴达摩院联合中国人民大学、中国科学院大学等机构发布行业内首个超导材料发现智能体 ElementsClaw，展示了一次完整的科学决策。其中，Elements 是一个 10 亿参数的大型原子模型；智能体 ElementsClaw 则作为整个科研流程的决策中枢，负责编排任务、调用工具并协调 4 个专用技能完成连续决策。

该智能体在 28 个 GPU 小时内，扫完 240 万个稳定晶体结构，从中识别出 6.8 万个超导候选材料，最终实验合成了 4 种，经实验验证全部具有超导性。这一科研智能体框架不仅能够应用于超导材料发现，还可迁移至固态电池电解质、多相催化剂、热电材料等不同材料研发场景。

更重要的是，4 种新型超导体并非来自同一种发现机制，而是代表了科研智能体开展科学探索的 4 条不同路径。

随着科研智能体逐步覆盖文献理解、假设生成、实验设计、工具构建与实验验证等环节，并与自动化实验平台深度融合，科学研究有望形成端到端的自主发现闭环。它的价值也将从提高单个环节的效率，演进为持续催生新知识的创新引擎。

ElementsClaw 构建 4 条科研探索路径

				
发现模式	结构模因引导	从头生成	结构重新诠释	数据库隐含生成
与既有知识的关系	• 基于已有结构规律 • 通过继承既有知识	• 完全脱离已有样本 • 此前不存在的新结构	• 修正既有知识 • 重新发现被长期忽视的材料	• 挖掘既有知识尚未显式发现的潜在规律 • 根据规律外推
发现材料	Zr_3ScRe_8	$HfZrRe_4$	Zr_4VRe_7	$Hf_{21}Re_{25}$

展望

科研智能体的能力扩展，依赖于强化学习阶段持续的反馈迭代，而真实世界中实验验证的效率决定了这一反馈循环的速度。相比互联网领域的任务，湿实验仍然周期长、成本高、自动化程度有限，是科研智能体扩展的重要瓶颈。未来，随着自动化实验平台、数字孪生等技术的发展，端到端的 AI 科研周期有望进一步缩短，推动形成更高频、更自主的科研迭代模式。与此同时，人类科学家与 AI 研究者也将共同定义问题，推动科学研究方法论发生新的变革。

AI 赋能绿色发展

AI for Sustainability

闲鱼循环经济
高精度风光分布图
小有可为“绿色赛道”
绿网计划生态守护

全球气候变化、生物多样性丧失与污染，构成相互关联的三重环境危机。应对这些挑战，不仅需要减少排放，也需要改变资源的生产、使用与废弃方式，保护生态系统及其承载的生命。

AI 为绿色转型提供了新的工具。通过识别复杂数据中的规律、改善预测与匹配、辅助监测和决策，AI 有望帮助人们更准确地认识环境变化，更有效地利用能源和物质资源，更及时地发现生态威胁。与此同时，技术的环境价值并不自动产生：它既需要控制自身的资源消耗，也需要与专业知识、治理机制和实际行动结合，将识别与分析能力转化为可验证的环境改善。

四条实践路径——从延长一件物品的使用寿命，到认识能源资源、守护野生物种，再到小有可为发动开发者参与环保——探索的是同一个问题：如何让 AI 能力回应具体的环境需求，并为更多主体参与绿色转型创造条件。

闲鱼

以 AI 支持闲置资源再流通

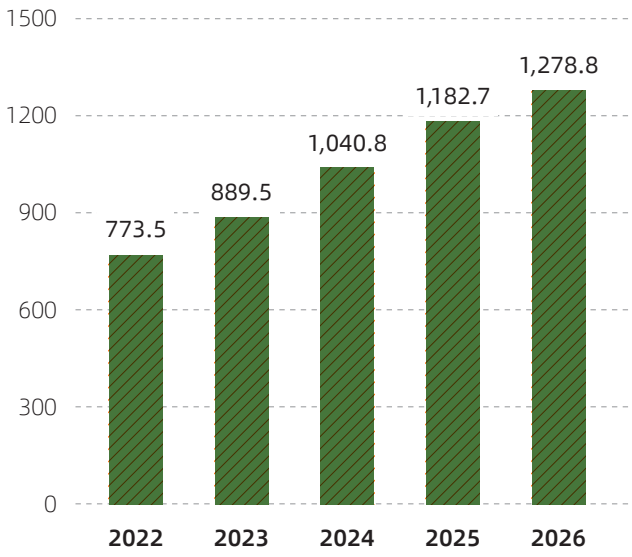
让一件仍可使用的物品找到下一位使用者，是循环经济贴近日常生活的方式。通过延长使用周期，并在一定程度上替代新品购买，闲置再利用有助于减少新增资源需求和废弃物产生。但从“家中闲置”到“重新使用”，往往还隔着信息整理、寻找买家、价格判断和交易沟通等环节。

数字平台为连接这些分散的供需提供了基础。一项对全球 246 个循环平台的研究指出，平台通过搜索与匹配、降低交易成本和扩大参与网络，支持旧物、可修复产品及材料重新流通，并将闲鱼列为跨品类循环平台的实例。

闲鱼作为中国领先闲置交易平台及生活方式社区，对闲鱼而言，平台的作用不仅是提供一个发布闲置物品的入口，也是在分散的个人之间建立信息和交易连接，让不同品类、不同成色的物品有机会继续发挥使用价值。

闲置物品回收利用避免碳排放

避免排放：万吨二氧化碳当量



闲置物品的非标准性带来的信息摩擦，是流通中的一个主要障碍。同一型号的商品，也可能因为成色、使用历史和配件不同而存在价格差异；卖家不一定知道如何描述，买家也常常需要反复询问。信息差使许多仍具使用价值的物品长期闲置，甚至还没来得及找到下一位使用者就被丢弃了。对于单件价值不高、却需要花费时间整理和沟通的闲置物品，降低这些成本尤为重要。AI 在这条链路上的着力点，是改善商品信息、供需匹配和交易响应，让更多仍具使用价值的物品更容易找到合适的使用者。

2026 年，闲鱼围绕供需两端分别部署 AI 智能体，尝试将分散的非标交易转化为可持续协助的服务流程，帮助用户更容易地发布、寻找和交易闲置物品。

供给侧，鱼卖卖利用多模态能力理解图片和文字，辅助识别商品品类、品牌和型号，生成标题与描述，并提示卖家补充成色、维修历史和配件等信息，供其核对完善。结合相似商品、历史交易价格和市场供需，估价功能为卖家提供价格参考，降低“不知道怎么写、也不知道卖多少钱”的操作负担。发布后，智能托管可协助处理重复咨询和履约提醒，减少用户持续跟进交易所需的时间。

需求侧，鱼买买尝试将买家日常、模糊的表达转化为具体的找货条件，理解用途、预算、型号和成色等要求，必要时进一步澄清，再从大量非标准商品中筛选候选项。AI 由此将平台的供需连接延伸到信息整理与沟通协助，支持用户完成原本较为繁琐的交易步骤。

截至 2026 年 10 月，闲鱼卖家智能体给出的建议采纳率 55%，商品动销率提升 6.5%。

经第三方审计，2026 财年（2025 年 4 月至 2026 年 3 月），按平台现有核算方法估算，闲鱼用户通过闲置交易与回收实现避免排放约 1278 万吨。这一数字反映平台整体循环活动的成效，AI 在其中的具体贡献尚难以单独量化。

随着 AI 逐步融入平台服务，我们希望让继续使用已有的物品成为更便利、更可信的选择。如果闲置消费只依赖节约资源的意愿，参与门槛仍然偏高；当交易体验本身足够顺畅，更多人会在日常中自然做出这个选择。

高精度风光分布图

AI+ 助力绿色能源发展

风电和光伏的规模越大，新能源之间如何协同的问题就越重要。风和光并不总是同时同地出现，通过跨区域调度，可以平抑发电波动、提高消纳能力。与吉瓦级智算中心的“算电协同”，又让这一问题变得更加迫切。但要实现高效的跨区域调度，前提是拥有一张覆盖全国真实设施的高精度地图。

要绘制这样一张全国尺度的高精度风光分布图并不容易。中国不同地区的地貌差异巨大，从沙漠、草地、雪地到森林、水域，对模型的泛化性要求高。阿里巴巴达摩院基于自研深度学习模型，结合覆盖全国、总量达 7.56 TB 的 0.5 米分辨率卫星遥感影像，首次高精度刻画出全国 31.9 万处光伏设施和 9.16 万台风机的空间分布。2026 年 5 月，该成果发表于《自然》杂志。

由此，分散的风光资源从宏观统计数据转变为可计算、可调度、可规划的绿色资源。研究显示，在统计期内（数据截至 2022 年），全国跨省协同可额外提升约 1000 亿千瓦时的绿电消纳能力，相当于全国风电和光伏总发电量的约 9.1%。

除了支持绿电的实时调度，这张数字底图也为储能配置、算力中心选址等前期资源规划提供了依据。

“小有可为”绿色赛道

关注与自然共生

2026“小有可为”AI 向善创新挑战赛，是阿里巴巴 ATH 事业群联合阿里巴巴公益、阿里云百炼共同发起的一项面向公众，征集公益问题 AI 解决方案的创新赛事，今年进入第二个年头，重点新增了绿色赛道。

第二届赛事共吸引 1111 支团队报名参赛，收到 AI 作品 558 件。在延续首届残健融合、普惠养老、乡村教育等场景的基础上，本届增设“绿色发展”场景，关注主题是如何“与自然共生”。

多款参赛 AI 应用聚焦自然声音识别，有的将动物叫声转化成音乐，有的通过声音揭示鸟类活动规律，有的用于记录亲子间的自然田野调查。这些应用背后，是阿里旗下开源生态的全模态模型 Qwen3.5-Omni、音频生成模型 ThinkSound 等。其中，ThinkSound 的训练筛选了 2532 小时的动物鸣叫、机械声和环境音等高质量音频样本。

此外，3 款“游园陪伴智能体”入围决赛，让动物园分散在公众号、饲养记录和讲解词中的科普内容得以走出园区，成为可随身携带的自然教育资源。

本届赛事还为参赛用户提供了 1 亿免费 Token，为这些 AI 应用的开发和试验提供基础支持。

绿网计划野生动物保护

让违规搜索变成自然教育

每年，全球非法野生动物贸易规模最高可达 230 亿美元，且交易正在不断向线上迁移。卖家通常使用暗号、变形词和多个账户，试图隐藏交易意图与轨迹。

以搜索“象牙”为例：一次看似无意的查询，背后可能关联着非法捕猎与贸易链条，而消费者往往并不了解其生态代价。平台掌握搜索、推荐、风控与交易等环节，具备消费者个体不具备的全链路干预能力。

绿网计划于 2020 年启动，以非正常搜索干预为核心行动。截至目前，绿网计划已覆盖超过 11 万个高风险屏蔽词，累计完成超 2 亿次非正常搜索干预。其中，野生动植物相关非正常搜索的拦截与转化累计超过 2500 万次。

2026 年，绿网计划从单一的非正常搜索干预，拓展为平台综合治理。依托 AI 能力，将传统“识别、拦截、干预”的治理方式，拓展到“援助、教育、协同”，覆盖延伸到生命守护、生态保护、成长守护和网络普法等领域。2026 年 6 月，绿网计划受邀参加英国皇家基金会“联合护野”商业论坛，分享中国互联网平台打击野生动物非法交易的治理实践。

当用户在淘宝等平台搜索涉及濒危物种的敏感内容时，系统不再只是返回“页面不存在”，而是推送法规解读、物种知识和保护倡议，将一次潜在的违规搜索转化为一次自然教育。如果消费者只是希望购买工艺藏品，平台也可以推荐合适的替代产品。

AI 扩展了治理的覆盖范围，尤其是理解更复杂的搜索表达，识别变形词、隐晦说法，以及隐藏在图片、视频中的潜在风险信息。AI 也被用于生成普法内容初稿、页面模板等，让平台能够更快针对不同物种补充相应的科普内容。绿网计划还向公众开放平台此类 AIGC 能力，鼓励公众主动参与内容创作。

在上述能力持续验证的基础上，2026 年，绿网计划发起“一亩田计划”，聚焦野生动物非法交易协同治理与替代性生态修复，将线上治理能力进一步延伸到生态修复、自然教育和公众参与。

绿网计划四大守护领域

生命守护

- 自杀干预
- 心情守护者计划

生态守护

- 动物驿站
- 一亩田

成长守护

- 青橙课堂
- 未成年人守护中心

公益普法与教育

- 非正常搜索干预
- 商品消费者提醒
- 网络普法

展望：

AI 本身既是能耗的增量来源，也是绿色转型的重要工具。如何让 AI 的能耗增长与其带来的减排效益相匹配，是这一代技术必须回答的问题。加速建设新型电力系统、让算力革命成为能源革命的一部分，正是关键的技术杠杆之一。

更长远来看，AI 正在改变绿色发展的实现方式：让过去难以观测的资源与生态变得可计算、可调度、可管理，也让更多个人和组织有能力参与其中。当技术进步与生态责任同向而行，绿色发展才具备持续的动力。

AI 赋能经济机会

AI for the Economy

魔搭社区
Y/OUR SPACE 亲橙空间
Accio 全球电商 AI 工作台
AI 乡村特派员

联合国将促进持久、包容和可持续经济增长，促进充分的生产性就业和人人获得体面工作，视为 2030 年可持续发展议程的关键目标之一。作为一项通用技术，AI 正在快速进入各行各业。扩大经济机会、促进包容性增长，是技术普惠的基本方向。

AI 技术的快速迭代与低成本扩散，正在降低创新与创业的门槛。更多个人、中小企业乃至传统产业不仅能够参与市场竞争，还能以更小的团队规模调用多种工具与能力，完成过去需要大量资源才能实现的复杂任务。

魔搭社区

让创新来自离问题最近的人

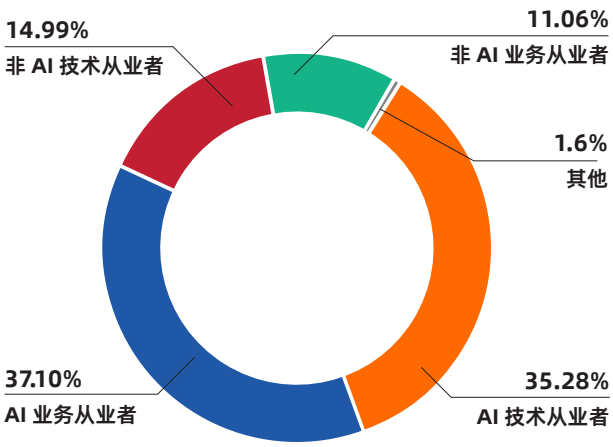
AI 正在扩展创新群体。开源模型、低代码开发工具和普惠算力不断降低创新门槛，让最了解真实问题的人，第一次有能力亲手解决自己的问题。不同行业的一线实践者，乃至残障人士等以往远离技术开发的群体，正从创新成果的接受者，转变为解决方案的创造者。

作为国内规模最大的 AI 开源社区，魔搭社区（ModelScope）正在成为这一进程的基础设施。截至 2026 年 9 月，社区已经汇聚 3300 万开发者，超过 25 万个开源模型、4 万个数据集向全球开发者免费开放。千问系列模型支持 201 种语言及方言，帮助不同文化背景的开发者开展本地化开发与部署。

2026 年，魔搭社区进一步向智能体（Agent）开发的基础设施演进。社区推出 Skills 中心，开发者可将复杂能力封装为可复用的技能模块，已逾 8 万个；上线 MCP 广场，汇聚逾 1.2 万个 MCP 服务并提供免费托管；同时全面开放模型、数据集等核心接口。这些能力共同降低了从模型调用到智能体构建的门槛。

AI 也在改变知识和经验的价值形成方式，让这些以往依附于个体存在的行业经验，沉淀为可复用、可迭代，甚至可交易的智能体或技能。社区发布的 AI 开源人才报告显示，

您当前学习 / 工作的领域是？【单选】



来源：2026AI 开源人才报告，P16

28% 的开发者并不具备专业 AI 背景，而是来自医疗、教育、制造等行业；越来越多离问题最近的人得以进入创新的现场。

这一点在社区衍生的公益实践中尤为具体。在魔搭社区参与发起的首届“小有可为”AI 向善创新挑战赛（2025 年）中，一位视障开发者针对通用视觉模型难以满足的出行需求，自主训练了面向视障群体的红绿灯识别模型；一位构音障碍开发者基于通义千问语音模型打造“燃言”应用，并

借助 AI 完成了人生第一次公开演讲。在他们身上，需求发现、能力开发与真实验证形成了完整闭环。

魔搭社区所做的，正是为这种闭环提供持续运转的土壤。魔搭社区向开发者开放模型、工具链和免费 GPU 算力，让开发者能以较低成本完成从 0 到 1 的创新探索；社区里 96% 的开发者明确愿意二次开源，又使这些成果能够不断被复现、迭代和扩散。

从学校到社区再到社会，部分项目逐步实现商业化或公益规模化落地，更多探索则不断沉淀为可复用的模型、工作流和智能体，以及 AI 原生人才，为下一轮创新提供基础。

Y/OUR SPACE 亲橙空间

超级个体线下社群

AI 浪潮下，创业的最小单元正在改变。大模型大幅降低了创业的技术门槛，超级个体、一人公司（OPC）等新创业形态涌现。他们既是灵活就业的新形态，也是社会创新中最为活跃的群体之一。

但是，在 AI 技术快速变化期，越是以个人为单位的创业者，越需要重新建立人与人、企业与企业之间的连接。他们既需要交流和学习最新的 AI 技术，也需要真实的业务场景、产业资源和协作伙伴，将 AI 能力转化为实际的创业能力。而这些连接，仍然主要在线下发生。

为此，2025 年 5 月，阿里巴巴不动产管理部推出创业社区 Y/OUR SPACE（亲橙空间），不是简单的租赁经济，而是兼顾对微型企业及 OPC 的支持。亲橙空间不仅提供了共享办公空间与一站式创业服务，更试图成为个体创业者与产业生态之间的“接口”：将阿里的大模型、算力等 AI 能力与电商等产业资源对接到普通创业者身边；引入魔搭社区及外部社群，发起技术交流与产业互动活动；对接当地政府 OPC 支持政策，推动算力、数据、场景、词元等要素补贴落地。通过这些助力最终帮助创新单元实现社会价值，并建设和扩大了共同繁荣的生态圈。

目前，亲橙空间租户约 60% 来自互联网公司前员工，典型的组织规模在 2-4 人之间。共享办公也因此不再只是传统业态的“共享空间”，而是在为这些超级个体重建原本由组织承担的连接、资源与协作网络。

来自亲橙空间的创业者，已经开始崭露头角。9 月，在首届 GOAI 世界人工智能开源大赛上，来自杭州云谷站的两名创业者，分别凭借“质链智能体”和“山猫巡游记”项目，从全球 91 个国家和地区的超过 1.4 万名开发者中脱颖而出，获得无界应用赛道的冠军和具身未来赛道的亚军。



Y/OUR SPACE 活动现场

截至 2026 年 9 月，从杭州西溪的首个站点起步，亲橙空间已在全国 24 座城市布局 31 个站点，累计服务 4000 余位创业者，社区稳定期入驻率保持 90% 以上，甚至有的社区出现“一席难求”。其中，与西湖区政府共建的魔搭杭州 AI 开发者中心（云谷 YS 站），已成为杭州人工智能开源生态的代表性社区。

Accio 全球电商 AI 工作台

降低中小企业参与全球贸易门槛

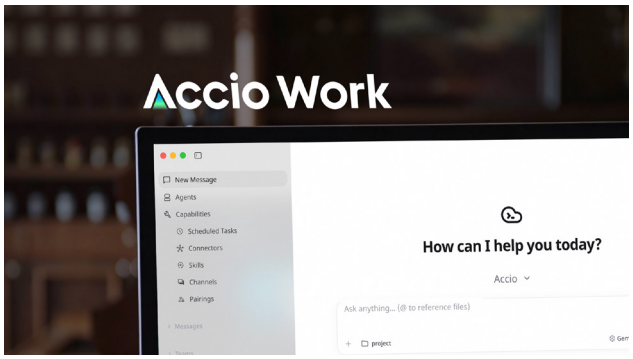
全球贸易的进入门槛一直不低。从市场调研、采购寻源、再到商品运营、跨境营销、客户服务等，整条链路需要专业团队协同完成，中小企业和个体创业者往往难以独立承担。随着全球贸易体系日益碎片化，小批量、多批次、快速响应成为新的竞争常态，跨境经营的门槛进一步提高。

2026 年，阿里国际推出面向全球贸易的“全球电商 AI 工作台”Accio，不仅打通淘宝、天猫、1688、抖音、拼多多、京东等多个国内电商平台，还打通了 Amazon、eBay、Shopify、Temu、TikTok 等多个海外电商平台。

依托 26 年国际贸易经验，它将出海关键商业流程沉淀为 100 多个可调用的技能，并通过多智能体协同，完成复杂业务执行。这些商业能力原本依赖组织分工才能实现。此外，阿里国际还开源了 CommerceAgentBench 测评基准，推动行业验证智能体在真实场景的表现。

这大幅降低了参与全球贸易的门槛。过去，搭建一家跨境网店需要数周时间和数万元投入，如今最快 30 分钟即可完成，词元成本仅需约 70 元左右；海外买家从商品搜索到形成采购意向的时间缩短了 80%，阿里国际站新增海外买家同比增长 40%。

截至 2026 年 7 月底，全球已有 1000 万中小企业使用



Accio Work

Accio 开展业务。与去年同期相比，阿里国际站上的中小企业词元消耗量翻了 10 倍；而 Accio 的词元成本，要比通用智能体低 50%。

随着商业能力不断技能化、智能体化，这些从人与组织中解耦的商业能力，有望进一步成为智能体之间协同与交易的基本单元。未来，越来越多个体将能围绕自身核心能力，组织和调用这些智能体，将全球贸易推往智能体协同的新阶段。

AI 乡村特派员

基层组织的 AI “脚手架”

AI 落地基层，难点往往不在技术本身，在于如何理解基层真实需求、融入组织流程，并形成持续运行的服务机制。这不仅要求 AI 适应基层，也要求基层组织完成面向 AI 时代的自我升级。

阿里巴巴自 2019 年起向欠发达县域派驻乡村特派员，协助当地对接产业和技术资源。在此基础上，2025 年 10 月进一步发起 “AI 乡村特派员” 计划，将这一机制从产业帮扶延伸到 AI 能力的基层落地。

特派员将基层的杂乱的问题翻译为产品需求，将地方长期积累的隐性知识翻译为模型能够理解的数字能力，将分散的 AI 能力翻译为能够持续运行的公共服务。这也是阿里 AI 能力触达基层的重要路径。

以浙江嵊泗为例，其船票体系、航线规则和天气变化都很复杂，长期以来只能靠本地经验来判断。外地游客是否自驾、上哪座岛，轮渡选择都不一样，获取信息的成本极高。特派员将这些隐性经验与数据系统，整合为 “AI 出行助手”，所有游客都能够调用这一公共能力。类似实践正在更多县域持续复制，推动 AI 不断理解地方表达、基层流程和真实场景。

截至 2026 年 9 月，“AI+ 县域” 智能服务平台已在 15 个地区上线，覆盖了 “政策咨询与营商环境”、“企业办事流程”、“12345 工单服务” 等十余个基层场景。帮助基层减少了一半的重复咨询的接待工作，投诉量明显下降。工作人员也得以 “解放”，专攻跨部门联办等复杂业务。

AI 乡村特派员如同一套 “组织脚手架”，帮助基层逐步形成持续吸收、沉淀和应用 AI 能力的机制。这激发了乡村和基层拥抱 “AI+” 的内生动力。在浙江嵊泗，30 个单位沉淀了 60 个应用场景。当这种能力真正留在地方，AI 向善便不再只是技术向社会的单向输出。

工单处理

时间缩短	投诉量	重复咨询
至 1-2 天	下降 30%	减少 50%

说明：相关数据来自部分 “AI+ 县域” 智能服务平台，并非全量统计数据，仅用于反映相关应用情况。

展望

AI 在扩大经济机会的同时，也在深刻改变工作内容和就业结构。大量重复性任务将逐步交由智能体完成，如何让技术带来的效率红利转化为更广泛的发展机会，而非加剧就业与能力的分化，是这一代技术必须回答的问题。其中的关键，在于企业能否从单纯部署 AI 工具，转向重构组织、流程和人才培养体系，帮助员工从重复劳动转向创造、协同与复杂决策，持续积累 AI 时代的人力资本。

更长远来看，AI 正在把创新、创业和全球贸易的门槛降到前所未有的低点，让更多个人和中小主体第一次有机会参与全球市场。但机会能否真正普惠，仍取决于模型、工具、平台和生态的持续开放。从电商时代降低交易门槛，到 AI 时代降低创造门槛，平台企业不仅是 AI 技术的发明者和应用者，也将成为 AI 生态的建设者和经济机会的创造者，让更多人拥有参与全球创新与商业的机会。

AI 赋能
安全包容的社区
AI for Community

绿网计划和千问 AI 心理陪伴
优酷无障碍剧场
淘宝 AI 视频通话
钉钉 A1 社工
高德“鹰眼守护”
“古壁生辉”壁画 AI 重现

AI 应该增加社区的关怀性。建设安全、包容、便利的人类住区与公共空间，是联合国 2030 年可持续发展议程的重要目标之一。这一目标也已从现实空间延伸到数字空间。但无论线上还是线下，公共服务仍面临覆盖不足、专业资源有限和响应不及时等问题。

AI 正在为这些服务提供新的技术基础。从人与数字空间的安全、陪伴与无障碍，到人与现实空间的社区服务和公共安全，再到人与历史空间的文化遗产保护，AI 正在将安全、便利与包容嵌入更多日常场景，让更多容易被忽视的需求获得被看见、被理解和被服务的机会。

绿网计划和千问 AI 心理陪伴

从拦截风险到缓解问题

心理健康已成为最具代表性的长尾公共服务需求之一。病耻感、经济压力以及专业资源分布不均等因素，让一些潜在的心理危机难以及时获得支持。更为关键的挑战在于，当求助意愿或轻生念头出现时，能否在第一时间得到回应。

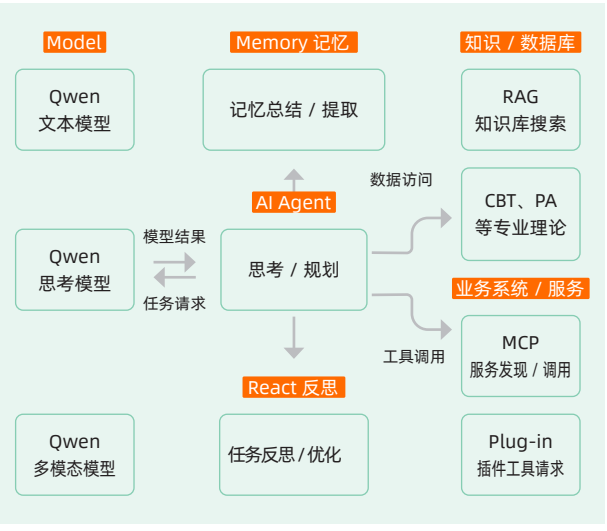
AI 仍然无法替代心理治疗，但可以让平台成为那个最先伸出援手的角色。基于 AI 陪伴与角色对话的技术积累，千问 AI 心理陪伴智能体与绿网计划展开合作，共同构建了一套从风险识别到情绪承接，再到专业转介的闭环支持体系。

2026 财年，绿网计划通过动态词库更新与 AI 语义识别，新增应规避关键词近 5 万个，包括抑郁类用品交易等非正常搜索。

当绿网识别到具有自杀、自残倾向等高风险搜索行为后，千问 AI 心理陪伴智能体将通过用户主动发起的对话，提供情绪疏导和行为建议，并通过持续的陪伴，帮助用户逐步稳定情绪；在必要时，还会引导专业人士介入。

相较商业场景，心理健康危机对安全性、可靠性和价值观一致性提出了更高的要求，也更容易触及模型的能力边界，推动陪伴智能体持续迭代。为此，团队联合华东师范大学等机构，将心理学量表与临床经验结构化为评估体系。针对极端风险场景，系统可引导用户进入线下心理救助体系，同时全流程严格遵循用户授权与合规要求。

截至 2026 年 8 月，通过绿网入口触发的心理干预体系累计产生约 109 万次对话交互，累计陪伴约 45.5 万人，单日峰值达到 1.6 万人次，并成功转介 2376 人。



优酷无障碍剧场

包容、便利的视听空间

随着信息获取、文化娱乐和公共服务不断向线上迁移，数字空间已经成为人们日常生活的重要公共空间。但如果一个空间无法被所有人平等使用，它的便利就仍然是有界限的。对于约 2870 万听障人士、1700 万视障人士以及老年群体而言，大量数字内容仍然存在获取障碍。

依托多模态大模型能力，阿里巴巴正在构建 AI 驱动的无障碍内容生产能力。AI 开始直接理解画面、对白、环境声音和剧情结构。在优酷无障碍剧场，AI 技术赋能内容适配，面向视障用户生成解说版，面向听障用户生成助听字幕版，在保持剧情完整性的同时，分别通过解说旁白和字幕标注，还原人物动作、空间关系、环境声效、情绪变化等，帮助视听障群体无障碍获取更多剧情信息。由此，用户体验不再只是停留在知晓发生了什么，更能理解它为什么发生。

AI 显著缩短了无障碍版本的制作周期，能够更快提供新热内容的无障碍版本上线；AI 解说语音包，也满足了更多残障人士个性化需求。从前，一部 2 小时的无障碍影视作品，从解说脚本撰写、旁白解说配音到制作上线，往往需要 1 个月的时间完成；如今借助 AI 技术，制作周期从原来的至少 1 个月，压缩至 3 天内。有了 AI 解说语音包后，视障人士还能使用胡歌等自己喜欢的明星声音，为自己全程讲述电影画面细节。

截至 2026 年 10 月，优酷无障碍剧场累计覆盖超过 9500 部（集）影视作品，解说版与字幕助听版持续更新，每月各上线 3 部电影、每周各上线 3 集剧集 。2026 财年，无障碍剧场累计播放超 121 万次。



优酷无障碍剧场志愿者走进校园

淘宝 AI 视频通话

让视障用户拥有更好的购物体验

对于视障群体而言，要真正融入数字生活，有赖于无障碍技术的持续发展。在淘宝，可识别的读屏人群达 151 万，其中 98 万为手淘月度活跃用户。这些用户不仅包括传统视障群体，也包括老花、低视力和各种其他原因导致的临时使用者。

从网页到移动互联网，淘宝在持续推进无障碍建设：读屏适配功能，帮助用户听取页面文字、识别操作控件；读光 OCR（光学字符识别）技术则能进一步读出图片中的文字。

但文字可读不等于商品可理解，颜色、款式和瑕疵位置等细节仍依赖视觉呈现。一个长期依赖视觉呈现和文字操作的电商系统，如何重新被设计，才能更好服务视障用户和读屏用户？

拍立淘 RTC（即淘宝 AI 视频通话），正通过 AI 技术提供一种新回答。用户打开摄像头，对准实物提问，系统通过图像识别、商品检索和语音交互，将“看”与“搜”衔接为连续对话。

并且，这项功能探索并非只服务于视障群体，而是淘宝将核心的商品理解与搜索能力，从视觉交互拓展至语音交互。当交互从“看与打字”扩展到“说与听”，潜在受益者还包括视力退化的老年人、照护孩子时双手被占用的家长，以及不熟悉文字输入的用户。

从读屏适配、读光 OCR 到拍立淘 RTC，淘宝持续探索从操作可达、信息可读走向自主判断。AI 技术普惠，不是要求所有人适应同一种界面，而是让同一种服务能够适应更多人。

钉钉 A1 社工助手

社工利器

社会服务高度依赖人与人的长期陪伴与信任。然而，一线社工、网格员和志愿者长期面临专业力量不足、人员流动性高的双重压力；更严峻的是，约 30%-50% 的工作时间被记录整理、材料归档、多平台填报等事务性工作占据。优质社会服务因此长期受制于人力供给，往往是人走了，经验也走了。

2026 年，阿里巴巴公益与钉钉携手，在复旦大学社会工作学系的专业支持下，将钉钉的 AI 软硬一体化能力转化为面向社工的赋能工具，并优先从儿童公益场景切入。

钉钉 A1 承载着钉钉 AI 的软硬一体化能力，通过语音转录、自动记录、结构化整理和辅助分析等功能，帮助社工完成记录、归档、填报等重复性工作。AI 不替代社工，而是把时间还给社工。走访、沟通、陪伴，这些 AI 无法替代、也最不该被压缩的能力，重新成为工作的主体。

当 AI 开始记录，服务过程便被完整留存。每一次走访、每一次干预、每一次回访，都不再止步于当下，而是逐步沉淀为可分析、可复用的实践经验；不同服务方式的效果，第一次拥有了比较与优化的基准。

AI 让机构获得了持续学习能力。尽管基层社会服务从业者流动性较大，但专业经验依托钉钉的组织平台持续沉淀，不再随个人流动而流失；新入职社工能够站在前人的经验上起步，这一能力的复制也不受地理边界的限制。

截至 2026 年 6 月，钉钉 A1 已服务超过 150 家公益机构，硬件一期申请覆盖 1000 名社工，首批 100 人已投入使用。同月，钉钉 A1 推出豆蔻医生版，将社会服务场景积累的方法论，进一步拓展至医疗等专业领域。

高德“鹰眼守护”

公共安全走向主动预警

道路交通事故仍是全球最主要的公共安全挑战之一，每年超过 100 万人死于车祸。传统交通治理主要集中于事故发生后的报警、救援和道路管控，而高速行驶中的风险往往发生在几秒钟之间。社会对交通安全的期待，开始前移到风险形成过程中的持续感知、实时判断与主动预警。

高德地图联合中国安全生产科学研究院研发了“鹰眼守护”预警系统。它能综合分析海量匿名脱敏的车辆速度、加速度与位置信息，结合空间智能和云端大模型，实时识别同向多车异常减速、停车等潜在风险，并通过地图和语音双通道向后方车辆发出秒级预警，为驾驶员争取“黄金反应时间”。2026 年，系统能力进一步扩展至冰雪、团雾等异常天气和特殊路段，覆盖更多公共安全场景。

安全的价值在于避免更多事故发生。截至 2026 年 9 月 30 日，鹰眼守护预警系统自上线以来累计播发预警已超过 713 亿次，日均预警次数达 1.9 亿次。据高德大数据，在 G2 京沪高速等试点路段，系统上线后，2025 年国庆期间日均万车事故数同比下降约 10%。

AI 也守护着阿尔茨海默老人的安全

2026 年，18 岁大学生董致澄基于阿里云搭建系统后端，接入千问大模型和高德地图，开发了协同寻人辅助系统“安归”。在“家属端”，千问能引导家属补齐走失时间、地点、衣着等关键信息；在“志愿者端”，高德可以辅助共享位置与标记已搜；最终“指挥端”由人汇总进展并调度任务。



左图山西平顺大云院严重褪色的壁画原样，右图是在 AI 助力下复现的壁画

“古壁生辉”壁画 AI 重现

用 AI 留住正在消失的文物

包容的边界，还可以进一步延伸到人类共同的历史空间。对于正在消失的文化遗产而言，最大的风险并不一定来自某一次突发事件，而可能来自时间本身。

山西平顺县大云院弥陀殿建成于公元 940 年，是国内现存少数五代木构建筑之一，被称为“五代建筑艺术的活化石”，也是全国重点文物保护单位。其中，弥陀殿东壁约 21 平方米的壁画《维摩诘经变》，是目前全国仅存、保存相对完整的五代寺观壁画。

但壁画正在消失。这不只是大云院的问题。中国现存壁画约 15 万平方米，许多都正在经历类似的命运。壁画保护需要高度专业化的知识，既要懂矿物颜料和传统工艺，也要理解不同时代的绘画风格。这样的人才全国屈指可数，到了县域和乡镇，很多珍贵壁画甚至难以获得基本的病害评估。

阿里巴巴公益联合西安美术学院、爱智岛、万相、堆友等合作伙伴，启动“古壁生辉”古代壁画 AI 重现工程，探索“专业技术人员 + 视觉大模型 + 智能体技能”的壁画数字重现方法。基于万相 Wan-2.7-Image 模型，利用 AI 对壁画识别并清理风化痕迹，推断残缺部分的原始构图，并在尊重原画画风的前提下，生成色彩饱满的数字重建图像。

西安美术学院设计艺术学院院长张犁介绍，第一阶段已经完成相关专业知识训练，“现在它已经是一个壁画专家”。目前这套技能已上线“堆友”平台，面向全国壁画从业者开放。

中国大量壁画散落在县乡小型寺庙，守护它们的未必是顶尖专家。当 AI 成为识别、评估和数字重现的工具，更多基层文保人员也就拥有了守护这些历史空间的能力。

展望：

AI 正在改变信息生产方式，也正在改变认知形成方式。未来，人们越来越多面对的是能够持续交流、不断反馈、主动适应个人偏好的智能体。特别是在心理健康、金融消费、公共舆论等高敏感场景，AI 形成的认知闭环，也可能不断强化用户已有偏见。安全与包容的数字能力的竞争，也将逐步从内容安全延伸至模型内生安全。

AI 能够降低知识获取与行动成本，同样也可能降低风险实现成本。当 AI 逐渐成为搜索、对话、购物和决策入口，它也开始影响用户相信什么、选择什么，以及最终采取什么行动。未来，数字空间安全与包容的重心将进一步前移，更可能以安全智能体的形式嵌入人机交互全过程。

第四章
面向未来的思考



进入社会，AI 再次演进的新起点

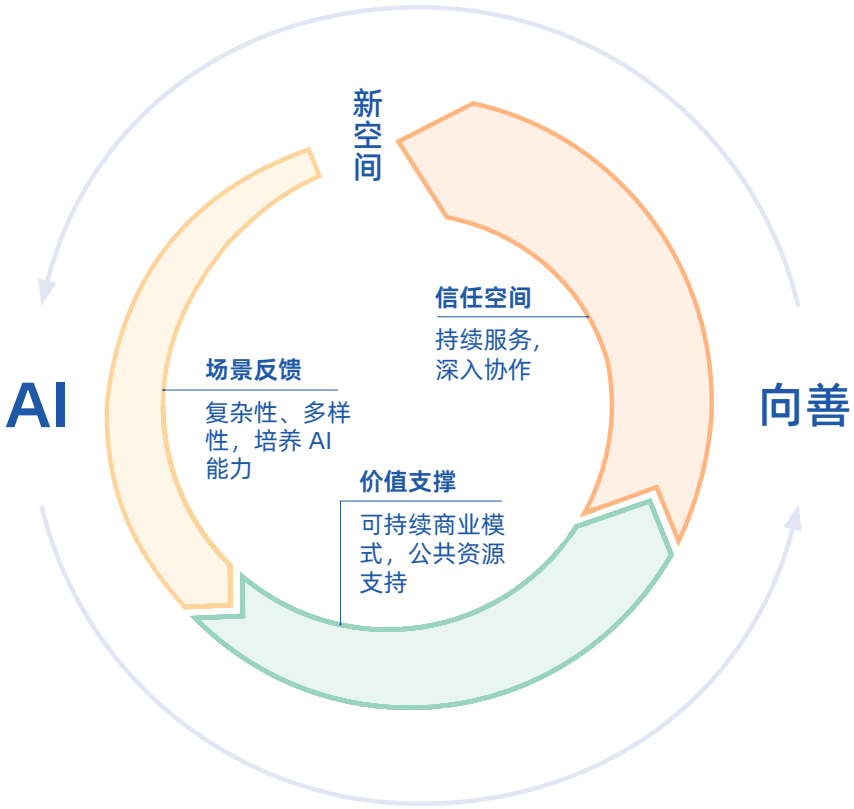
今年行动报告中的 26 个案例，全部基于真实的社会需求。对 AI 技术发展来说，最好的土壤就是真实的场景；在实验室里被发明出来的技术，必然会经历一个在具体社会场景中被“再发明”的过程；这个过程难免需要试错、离不开反馈和调整，而这些努力恰恰又构成技术演进的全新养分。

因此，向善不仅是我们发展 AI 的意义之所在，也是促进 AI 进一步发展的结构性动力。在去年的报告中，我们就提出“AI 与向善互相促进”（AI for Good is Good for AI），要在向善与 AI 发展之间构建一个相互强化的飞轮，需要三个回路相互支撑：

回路一：场景反馈。AI 进入真实而困难的场景，如帮医生读片、帮乡村教师备课、帮视障人士购物，可以获得实验室里得不到的反馈。向善场景的复杂性和多样性，本身就是训练更强 AI 的稀缺资源，培养 AI 系统行动的能力。

回路二：价值支撑。AI 向善项目解决真正重要的问题，需要持续的投入，而在能力的积累与提升过程中，有可能转化为可持续的商业模式或获得公共资源支持。这一环节要求每一次循环都会带来一定程度的提升，因而也非常具有挑战性。

回路三：信任空间。当 AI 可以令人放心地从单点试验走向持续服务，进入此前难以进入、却具有重要社会价值的领域，社会给予它的信任和空间随之增加，有机会从局部辅助走向更深入的协作。



每一次真实场景的应用落地，都是一次 AI 与社会的真实对话。对于 AI 企业来说，只有通过这样的飞轮，大量分散的场景探索，才有可能汇聚为系统性的能力积累和制度学习，成为智能再次演进的必要条件。

求真与向善：面向未来的八个问题

过去一年，人们感叹于变化是如此之快、如此偏离我们所习惯的判断。这提醒我们要时刻保持谦逊。在 AI 时代，我们面临一个共同的难题：塑造技术方向的最佳窗口期，恰恰在人们理解相对最不充分的时候。具体来说，在技术发展最初的快速演进期，很难预测其社会影响；而一旦技术成熟并嵌入社会后，想要控制其负面后果或做出改变就变得非常困难且成本高昂。[1]

联合国人工智能问题独立国际科学小组发布的首份报告也提出，政策制定者面临证据困境：要么在科学依据不足的情况下立即做出影响重大的人工智能治理决策，要么等待证据齐备，但那时再进行干预或许为时已晚。

因此，当我们在热议 AI“向善”的同时，我们不能放弃“求真”问题的探讨，也就是要不断地基于最新的事实来做调整。如果不去创造和使用 AI，就无法真正理解 AI；如果不充分地讨论和理解 AI 和社会的关系，就难以构建敏捷治理来解决发展中的问题。

在讨论和准备这份报告的过程中，我们也在不断提出“向善”的“真问题”，这既是一家 AI 企业的内省，也是和社会充分讨论的期待：

1. 效果验证：如何判断 AI 带来了真实、持续的改善？

如何区分技术性能、服务效率与最终受益，通过适合场景的评估识别 AI 的实际贡献？如何关注长期效果、群体差异及潜在不利影响，而不只统计使用规模？

2. 规模化安全：如何更早发现风险，并限制其影响？

智能体开始执行任务后，如何衔接发布前评测与运行监测，在必要时暂停、回退或转交人工？如何履行适用披露义务，并进一步说明评测范围和局限，同时避免泄露可被滥用的细节？

3. 开放与治理：如何选择与模型能力相适应的开放方式？

如何结合模型能力、预期用途及可预见的误用，评估开放范围、发布条件和配套措施？如何借助外部评测与社区反馈改进后续发布决策，既支持创新，也不将开放直接等同于安全？

4. 资源与环境：如何让应用增长与环境目标相协调？

如何结合效率优化、清洁能源、水资源管理和任务调度，同时关注单位效率与绝对资源消耗？如何明确计量边界，在适用要求和数据条件下完善训练、推理及基础设施环境信息披露？

5. 责任与协作：多方参与的 AI 服务，如何形成清晰的责任链？

无论模型是否开放，如何明确服务各参与方的授权、监督、告知和事件响应职责？发生问题时，如何通过必要记录、调查、纠正及用户救济落实责任，避免复杂技术链条模糊责任边界？

6. 人的能力与自主性：如何让 AI 辅助支持人的长期发展？

如何保留教育和专业服务中必要的思考、练习与核验，并观察使用者离开 AI 后的能力？在陪伴和个性化服务中，如何通过用户控制、明确告知及人工支持，减少过度依赖、维护自主选择？

7. 职业过渡：工作内容改变后，如何支持新的成长路径？

如何观察不同岗位的任务与技能变化，而不将效率提升直接等同于就业增减？如何通过岗位设计、导师支持、在岗学习及转岗培训，为新人和受影响群体保留成长阶梯？

8. 普惠可及：如何从“能够获得”走向“有效使用”？

如何识别语言、设备、成本及无障碍适配等实际障碍，让服务对象参与设计和评估？如何将资源支持与培训、本地服务及维护结合，验证投入是否真正改善受益机会？

这些问题没有单一的技术答案，也难以由一家企业独立解决。推进相关探索，需要在遵守法律法规、保护基本权益和满足必要安全要求的前提下，将实际效果、环境影响与普惠目标纳入技术和业务决策。以可验证的证据改进实践，以明确的责任支持协作，是我们理解 AI 求真与向善的重要方向。

[1] 这一现象在技术哲学中被称为控制困境（Collingridge dilemma）。参见 Collingridge, D. (1980). The social control of technology. Frances Pinter.

主要参考文献

1. 《人工智能问题独立国际科学小组的初步报告》（Preliminary Report of the Independent International Scientific Panel on AI），2026，联合国
2. 《促进平台经济大中小企业协同发展行动方案（2026–2028 年）》，工信部等七部门
3. 《智能体规范应用与创新发展实施意见》，网信办
4. 《守己利他：智能时代做负责任的技术》，阿里巴巴，上海人工智能实验室
5. 《2026 阿里巴巴环境、社会和治理（ESG）报告》，阿里巴巴
6. 《阿里巴巴 AI 向善行动报告 2025》，阿里巴巴
7. 《2026AI 开源人才报告》，魔搭社区
8. 阿里巴巴研究院系列研究报告
9. 《技术革命与金融资本》，卡罗塔·佩雷斯（Carlota Perez）
10. The Social Control of Technology, David Collingridge
11. Resilience in the AI Era, OpenAI
12. Anthropic's Advanced AI Framework, Anthropic
13. An Economy of AI Agents, Gillian K. Hadfield, Andrew Koh
14. A.I. and Our Economic Future, Chad Jones
15. Funding AI for Good: A Call for Meaningful Engagement, Hongjin Lin, et al.
16. A definition, benchmark and database of AI for social good initiatives, Josh Cowls, et al.
17. Agentic AI and the next intelligence explosion, James Evans , et al.
18. Advancing solar and wind penetration in China through energy complementarity, Hu, Y., Jiang, H., Zhang, C. et al.
19. Colorectal cancer detection using noncontrast CT and deep learning: a multicenter and international cohort study, Chen X, Qiu M, Zhang J
20. Harness Engineering for Self-Improvement, Lilian Weng
21. Ask an AI expert: What exactly is the full stack? Molly McHugh-Johnson, Google
22. Learned Hardware/Software Co-Design of Neural Accelerators, Zhang Shi et al.
23. The health of nations: Stronger health, stronger economies, McKinsey Health Institute
24. Cancer burden and control in China: landscape, trends and challenges, Wanqing Chen, Chen Wu, et al.
25. MM-StoryAgent: Immersive Narrated Storybook Video Generation with a Multi-Agent Paradigm across Text, Image and Audio, Mengyue Wu, Ming Yan
26. The future of artificial intelligence and the mathematical and physical sciences , NSF, Andrew Ferguson, et al.
27. Agentic Fusion of Large Atomic and Language Models to Accelerate Superconductor Discovery, Wenbing Huang, et al.
28. Oyster-II: Reinforcement Learning for Constructive Safety Alignment in Large Language Models, Alibaba AAIG
29. State of Open Models: Summer 2026 Observations. HuggingFace
30. Reasoning Models Generate Societies of Thought, Junsol Kim et al.
31. The Economics of Recursive Self-Improvement, Tom Cunningham et al.


Alibaba