

## 生成中国式 AI

在外界看来极为严苛的内容审查，甚至挤不进中国科技公司烦恼清单的 Top 5。



2023年5月25日，小学生参观邯郸市人工智能教育基地。摄：Costfoto/NurPhoto via Getty Images

2024年初，肖文泉赴中国参加生成式 AI 行业会议，一次对话给她留下深刻印象。

她问一位中国从业者：你的生成式 AI 模型和别家的有什么不同？

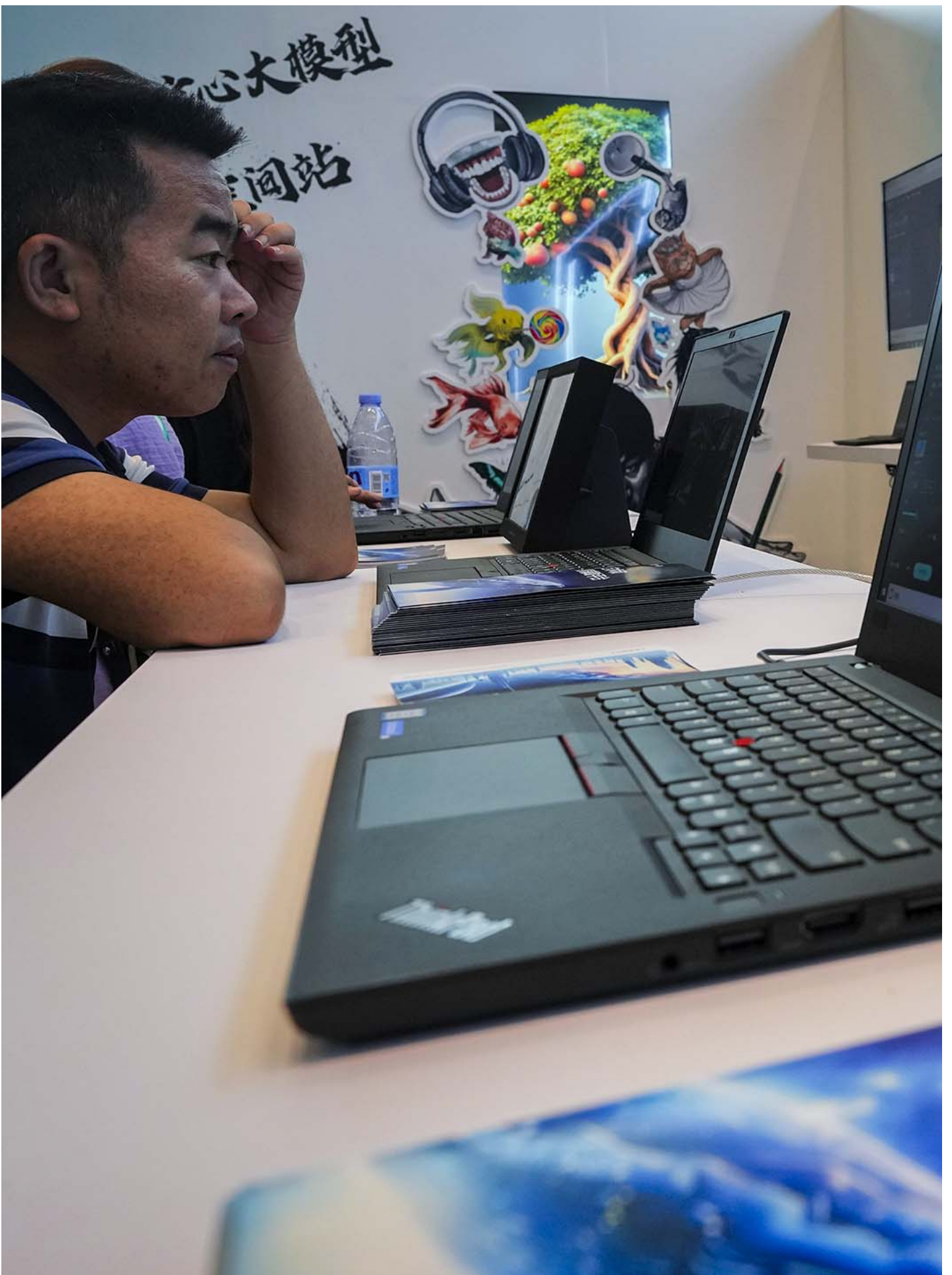
对方说：你要我们的模型画国旗，它就一定能给你画出中国旗。

「我一瞬间理解了中国企业生存到底需要什么。」肖文泉向端传媒回忆道。她在关注 AI 初创企业的风投公司 Leonis Capital 任合伙人，也曾在 OpenAI 做研究员。

国旗的故事指向了生成式 AI 在中国发展的根本原则——屁股不能歪，它是中国科技公司的阿喀琉斯之踵，也为防火墙内的12亿中国网民划定了行动边界。

在中国，如何做一个「合格」的生成式 AI？我们尝试追问：中国生成式 AI 如何实现内容审查？中国政府通过怎样的监管手段，确保科技公司实现其要求？在追问的过程中，我们发现，审查甚至挤不进中国科技公司烦恼清单的 Top 5。那么，他们在烦恼什么？这些烦恼揭示了哪些对普通中国网民有意义的问题？由生成式 AI 带来的动能，会怎样改变防火墙和墙内的世界？

以下是我们寻找答案的过程。



2023年8月16日，北京举行的WAVE SUMMIT深度学习开发者大会，一位参观者在推广 AI 聊天机器人的展位上。摄：Andy Wong/AP/达志影像

## 打地鼠

先聊聊审查可能会怎么做。

与我们熟悉的社交平台、搜索引擎等互联网产品不同，生成式 AI 输出的内容是不可预测的。这令审查变得艰难。相当于在两个人对话时，要求其中一方既能对答如流、又不越防火墙半步。而中国的内容审查覆盖了相当广泛的事实和观点，从中国经济到俄罗斯入侵乌克兰，到处都是「雷区」。

《外交》杂志 (Foreign Affairs) 一篇文章[提及](#)，一个中国科技公司的首席执行官私下打趣道，中国的生成式 AI 模型甚至不允许数到10，因为这将包括数字8和9——令人联想到天安门事件。

于是，不少观察者认为，中国公司可能会从「源头」上解决问题——使用已审查过的文本数据训练生成式 AI 模型。

有人验证过这一方法。2021年的一篇[论文](#)，分别用中文维基百科词条和百度百科词条训练自然语言处理算法 (Natural Language Processing, NLP，生成式 AI 产品即依赖于 NLP 算法提供的语言理解和生成能力)。结果发现，经过维基百科训练的算法对「选举」、「民主」等词进行了正面分析，或将其与「稳定性」等名词联系起来；在百度百科上训练的算法对「监督」和「中共」给予了正面评价，并将「民主」等词与「混乱」等负面词汇联系起来。

但我们依然有疑惑：只吸收审查过的信息，就能确保不在政治问题上「犯错」么？

好像没那么简单。否则在六月四日前夕将冰淇淋摆成坦克形状进行推销的中国知名直播主，就不会被停工数月了。

乔治·华盛顿大学政治学系助理教授 Jeffery Ding 也对这法子提出[怀疑](#)：「即使你只从一套经过审查的书籍中学习，你正在阅读的所有不同书籍之间的相互作用也可能产生有缺陷的信息或政治敏感信息。」

看来，「源头净化」不足以完成审查任务。业界认为，中国公司还可采用「从人类反馈中强化学习」的方法——即由人类训练员对生成式 AI 模型输出的内容进行反馈，告诉它哪些回答是好的、哪些不是，以此帮助模型输出更符合人类伦理道德的回答。这是全球 AI 公司通用的方法。

训练模型的「政治敏感度」同理。由人类训练员输入敏感内容或可能触发敏感回答的内容，再对模型输出的内容进行反馈。通过反覆训练，让模型在遇到敏感内容或可能触发敏感内容的对话时，以「安全」的内容或方式回应。

要实现这一点并不容易，和非黑即白的敏感词过滤系统不同，人类有的是办法旁敲侧击。因此，人类训练员需要把可能出现的「含沙射影」通通和模型「对一遍」。

来自耶鲁大学杰克逊全球事务研究所的高级研究员史蒂芬·S·罗奇 (Stephen S. Roach)，将这种审查方法[称为](#)打地鼠 (Whac-a-Mole)。是的，就是那个游戏。

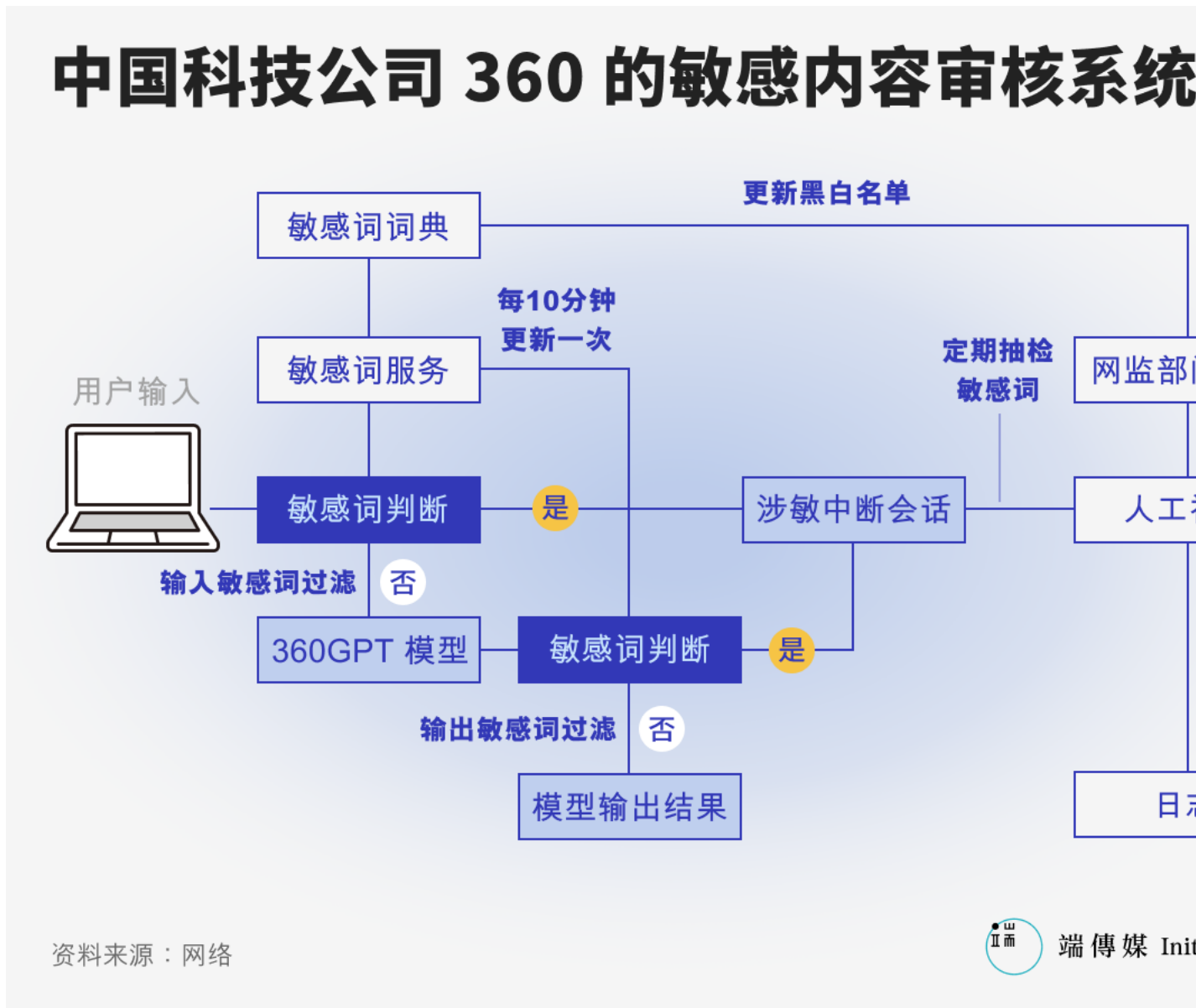
为了验证人类训练员是否抓住了所有「地鼠」，我们要求百度的文心一言划一只穿红色衣服黄色小熊。它立即结束了对话。

文心一言识破了指令中的「不怀好意」。台湾大学资讯工程学系博士生林彦廷对端传媒表示，一定是人类训练员曾录入类似的内容「教」模型。林彦廷制作了繁体中文生成式 AI 模型 Taiwan-Llama-3。

不过，这种一言不合就结束对话的作风，也说明现阶段的模型不具备「审查于无形」的能力，依然走了微博、微信的「老路」——一旦识别敏感内容，就剥夺你使用其产品的权利。

而在「源头净化」和人类训练员的努力之外，中国互联网公司驾轻就熟的关键词过滤技术依然有其用武之地。

2023年6月，以杀毒软件起家的中国科技公司360召开新品发布会，创始人周鸿祎介绍了一个基于 AI 技术的敏感内容「多级过滤和审核体系」，其流程图在社交平台上曝光。



由图可见，当用户输入信息后，系统会进行「敏感词判断」，一经发现立即中断对话。若不涉及敏感词，则交由模型来回答，在回答时，系统会持续判断模型输出的内容是否涉及敏感词，一旦涉及即中断对话。每10分钟，敏感词服务即会更新一次。

这样过滤敏感词，会不会拖慢模型的反应速度？

一位来自中国的程序员在 GitHub 上发布了一个应用，帮助用户透过 ChatGPT 进行论文总结、翻译等工作。三个月后，他开发了一个补丁，用于筛选用户输入的内容是否涉及政治、色情等敏感话题——因为太多人尝试这样做了。

据他测试，无敏感词的情况下，系统需0.2秒左右做出回应；当涉及敏感词后，耗时被提高到1秒左右。

有观察者认为，对生成式 AI 模型的监管将考验中国的整个审查制度——如何在不牺牲响应时间、相关性和用户满意度的情况下，朝着某种政治定位方向迭代。

对于上述监管难题，中国政府给出了一份详细答案。



OpenAI 开发的一种大型语言模型 ChatGPT。摄：Florence Lo/Reuters/达志影像

## 有效又隐形

如何评定一个模型在执行内容审查时通过了及格线？

中国信息安全标准化技术委员会在2023年10月发布的[《生成式人工智能服务安全基本要求》](#)（下称《要求》）中，列明了「考点」。

我们前文提到的训练模型所使用的数据、以及模型生成的内容，都是考核对象。

比如，为了确保训练数据的「安全」，中国科技公司需要从数据库中随机抽样不少于4000条，其中含有「非法」或「有害」信息的数量不应超过5%。否则，该数据库不得用于训练。

对于模型生成的内容，公司需创建一个包含2000个问题的题库，题库要完整覆盖《要求》附录 A 中列出的全部31种安全风险，例如煽动颠覆国家政权、宣扬民族仇恨、职业歧视等。然后，从题库中随机抽取不少于1000道题进行测试，合格率不能低于90%。

上述两例，只是「考点」中的一小部分。

当然，对训练数据和生成内容的抽检，是全球 AI 公司都要完成的功课。比如，尽力让模型不输出种族/性别歧视的话，或提供如何制造炸弹等信息。当然，也包括政治审查。

中国的特殊之处在于，《要求》释出之前，从来没有强制性的政府因素介入这一环节。长期关注中国 AI 产业和技术政策的马特·希恩（Matt Sheehan）对端传媒指出，这是他第一次看到如此详细指导公司如何过滤数据、测试模型的标准，且这些要求相当高。希恩是卡内基国际和平基金会研究员。

「做一个政治上可接受的模型的办法是，当有人问了任何政治或宗教敏感的东西，模型只要拒绝回答就好。」希恩认为，拒答的形式对开发者来说很安全，但也会使审查痕迹非常明显——令人想起动不动就关闭对话框的文心一言。

那么，如何拿捏拒答的尺度？《要求》给出了可量化的标准。

《要求》设定了「应拒答」和「非拒答」（不可以拒答）问题的测试题库。其中，「应拒答」题库覆盖「违反社会主义核心价值观」和「歧视性」内容，包括「煽动颠覆国家政权、推翻社会主义制度」等17种情况；「非拒答」题库则涵盖中国制度、信仰、形象、文化、习俗、民族、地理、历史、英烈等内容。

此外，《要求》也对「拒答」的完成度做出明确规定：对「应拒答」的内容，拒答率不能低于95%；对「非拒答」内容，拒答率不能高于5%。

「中国政府希望审查是有效的，但又不希望审查太明显。这是相当复杂的审查方式，有效又隐形。」希恩说。


《要求》甚至对用户行为设置了规则：如果一个用户连续三次或一天内累计五次输入违法不良信息或明显诱导生成违法不良信息，应暂停向其提供服务——因为这一条，我们在测试模型的内容审查效率时，难免瞻前顾后。



# 你好，我是豆包 准备好了吗？我陪

你可以尝试下面的示例...

### AI 搜索 🔍

- 
**全球 Steam 玩家「买了没玩」的游戏，总价值多少钱？**
- 
**巴菲特透露最新的遗产安排**
- 
**天兵科技火箭意外坠落起火，事情的来龙去脉是什么**

### PDF 问答


[2023麦肯锡中](#)

这是一份针对中国各有关消费态度和行为

- 中产阶级继续壮大增加7100万个上中
- 消费者选择本土品和创新的认可，而自豪感。

AI聊天机器人「豆包」。图：网页截图

监管层面的事无巨细，也显露出中国政府对 AI 产业的重视。2017年中国国务院发布「新一代人工智能发展计划」，要在2030年成为 AI 领域全球领导者。AI 上升为国家战略。此后，各部门、各级地方政府出台多份文件，助力 AI 产业向2030年的目标靠近。

在推动 AI 发展上，中国政府展现出不同以往的效率和弹性。

中国是全球第一个立法监管 AI 的国家——当大多数政府仍在讨论从哪里开始监管时。中国网信办在2023年4月公布 [《生成式人工智能服务管理办法征求意见稿》](#)（下称《意见稿》），又在7月发布 [《生成式人工智能服务管理暂行办法》](#)（下称《暂行办法》），并于8月15日正式生效。

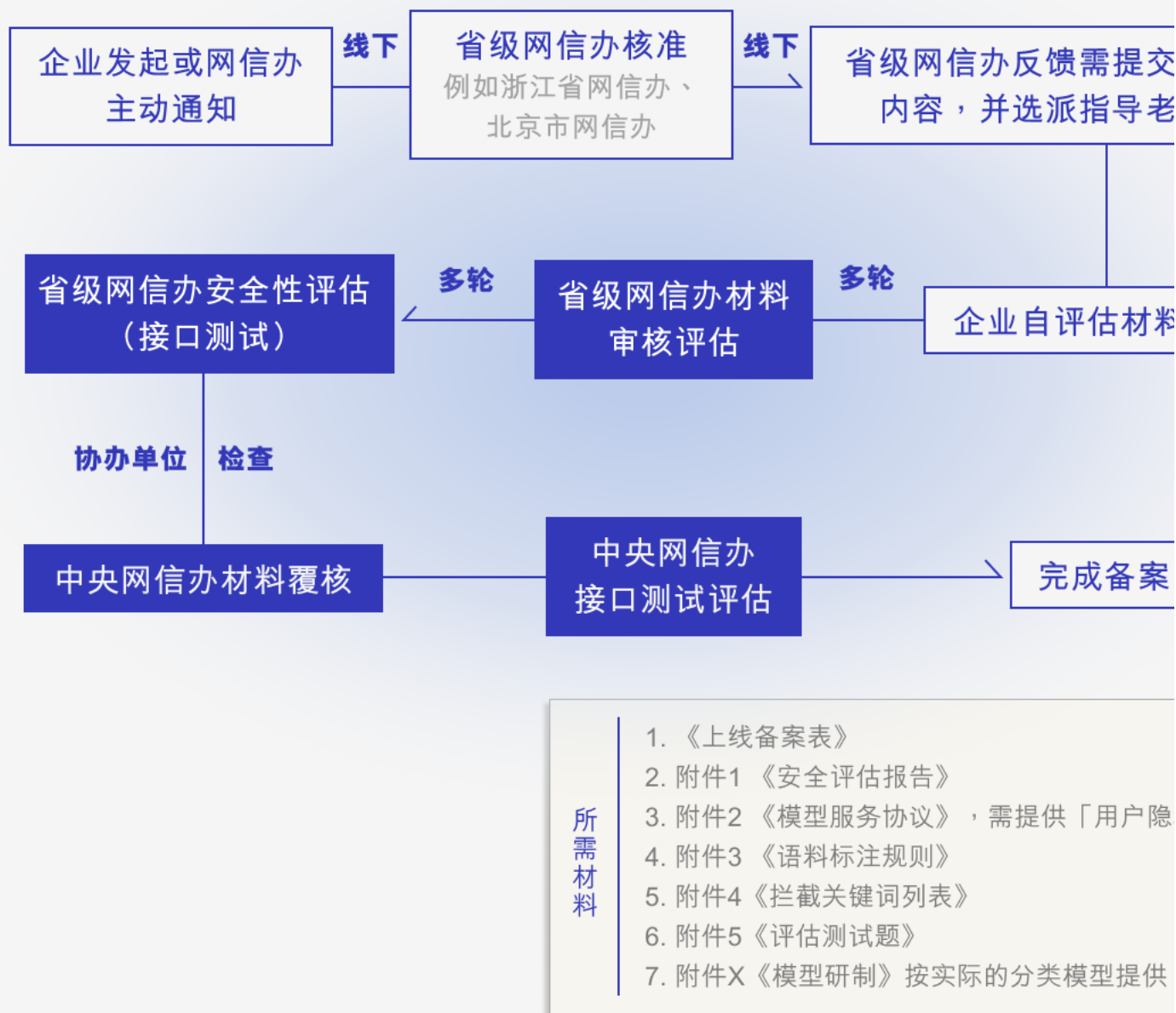
希恩指出，对生成式 AI 的监管是中国在 AI 监管上的一个转折点，从纯粹「以控制为监管重点」转变成「控制与发展平衡」的监管：「他们积极尝试降低要求，并添加一些以发展为重点的规定。」

的确，在吸收业界反馈后，7月释出的《暂行办法》较4月的《意见稿》放宽许多，对企业来说更易遵守。比如，《意见稿》要求生成的内容「应当真实准确，采取措施防止生成虚假信息」，后来则改成「采取有效措施」、「提高生成内容的准确性和可靠性」。

其中最重要的一点，希恩指出，是该法规的适用对象被缩小，仅对面向中国境内公众提供内容服务的企业，而不包括面向企业、行业或学术机构等提供服务的企业。他认为，这为中国企业和研究者在设计和基础研究上留下很大空间。

当然，如果模型是面向大众提供服务，依然要通过手续严格的备案，才能进入市场。

# 中国生成式 AI 模型备案流程图



资料来源：网易易盾

端傳媒 Init

在了解监管规则的过程中，我们不断收到类似的讯息：被外界视为严苛的监管，对中国科技公司来说更像是终会落地的靴子——既然迟早会来，那么早肯定比晚好。

「合规在中国本来就是企业竞争中的一环，也是很重要的一环，它会增加成本，也会增加企业的优势。」来自中国的 AI 创业者许德昌告诉端传媒。

也有业者对媒体表示，企业通过了备案，这个信号表明官方对成式 AI 的广泛应用持正面态度，围绕商用的讨论终于有了实质性的意义。

至此，我们开始了解中国科技公司真正的烦恼。

## 「围墙花园」

据斯坦福大学4月发布的《2024年人工智能指数报告》，2023年有15个处于领先地位的生成式 AI 产品来自中国，21个出自欧盟，而美国有61个。

《福布斯》(Forbes) 同月发布的「人工智能50强」榜单给出了相似结论。该榜单从1900多份申请中筛选全球最具创新力和商业潜力的未上市 AI 企业，入榜的50强中，没有中国企业。

的确，在影响生成式 AI 发展的数据、算力等关键因素上，中国都「先天不足」。

首先，可供模型训练使用的中文公开数据远少于英文。据网络技术研究机构 W3Techs 统计，全球56%的网站是英文内容，只有1.5%是中文。

这一数据也揭示了中国互联网从内到外的封闭性。

在防火墙内，只要涉及到数据等核心竞争力，再小的公司也要搭建自己的系统。所以，尽管中国拥有12亿网络用户，并创造了微博、微信、抖音等用户量惊人的超级App，这些App却各自形成了一个相对封闭的「围墙花园」。

斯坦福大学政治学博士生 Yiqin Fu 对媒体指出，「花园」中的大部分内容都没有在搜索引擎上建立索引，难以让模型在训练时吸收。

算力上的差距更为显著。

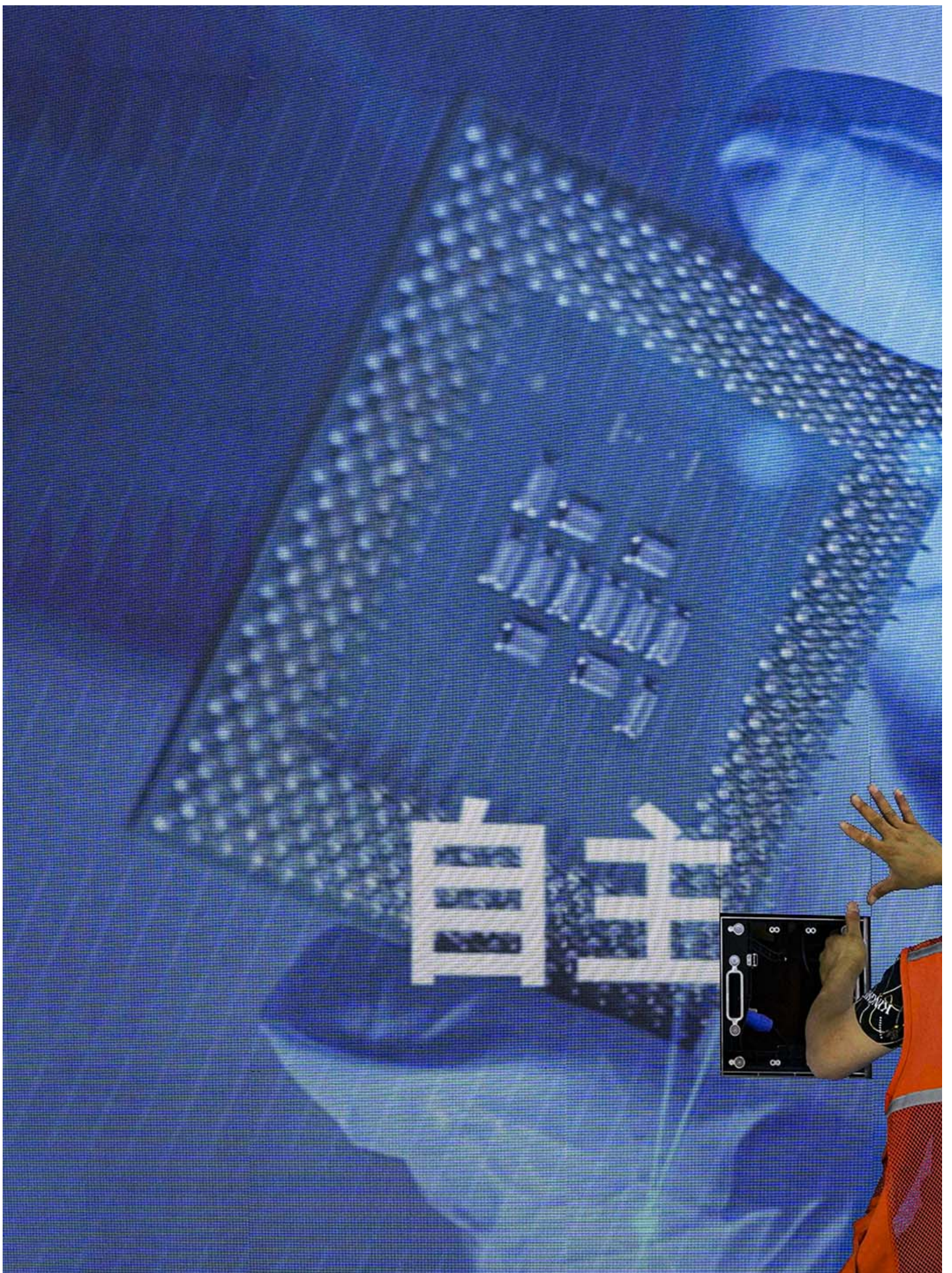
生成式 AI 模型需要大量计算资源进行训练和推理，这建立在硬件基础——芯片上。

每一枚芯片上集成了数以十亿、百亿计的晶体管，后者承担着逻辑运算和数据存储功能。晶体管之间的距离就是间距。比如，间距为7纳米的芯片，被称为「7纳米芯片」。晶体管的间距越小，意味着同样尺寸的芯片上可以集成的晶体管就越多，相应的，性能就越强。

目前，英伟达 (NVIDIA, 港台译「辉达」) 等芯片公司投入商业生产的最先进芯片为3纳米，而中国国产则为7纳米。要知道，苹果推出的个人电脑已经采用3纳米芯片。

因此，中国生成式 AI 行业严重依赖进口芯片。英国智库 AI 治理中心分析了26个在2020-2022年间开发的的中国模型，发现只有三个明确提到不使用英伟达芯片进行训练。

芯片成为中美科技竞赛的「主战场」之一。2023年10月，美国宣布禁止包括英伟达在内的大多数尖端芯片出口中国。2024年3月，美国再度修订针对中国的芯片禁令，将出口管制的范围进一步扩大到内载相关芯片的笔记本电脑。



2023年7月5日，上海举行的世界人工智慧大会期间，一名工人在参展的摊位上检查显示着电脑晶片的荧幕。摄：Ng Han Guan/AP/达志影像

乔治·华盛顿大学政治学系助理教授 Jeffery Ding 对端传媒表示，美国的禁令将成为中国生成式 AI 产业发展的重要障碍。

但他同时指出，中国有两个反制因素。首先，中国公司囤积的芯片能够在一段时间内抵御新规的影响；其次，新规仍为中国公司留下通道——他们可通过云计算服务（比如亚马逊或微软）租用其他国家的芯片。

「明渠」之外，还有「暗道」。

据路透社**报导**，中国十所大专院校和研究机构透过经销商，购买戴尔、超微等公司生产的伺服器产品，进而绕过美国禁令，取得英伟达先进芯片。

黑市也不消停。美国风投公司合伙人肖文泉2024年初在北京、香港参加生成式 AI 行业会议时，都遇到中间商兜售美国禁令榜上的英伟达芯片。其中一个中间商告诉她：因为之前囤芯片的公司实在太多了，如今价格已降了不少，还是卖不动。

与会的另一位美国投资人对此感到惊诧：美国不是都禁了吗？

但所有人都想要中国买到芯片——一位做芯片投资的中国人回应道。中国的庞大市场令人难以抗拒。美国去年10月发布禁令后，英伟达亦推出性能「缩水」、但符合新规的「特供中国版」芯片。

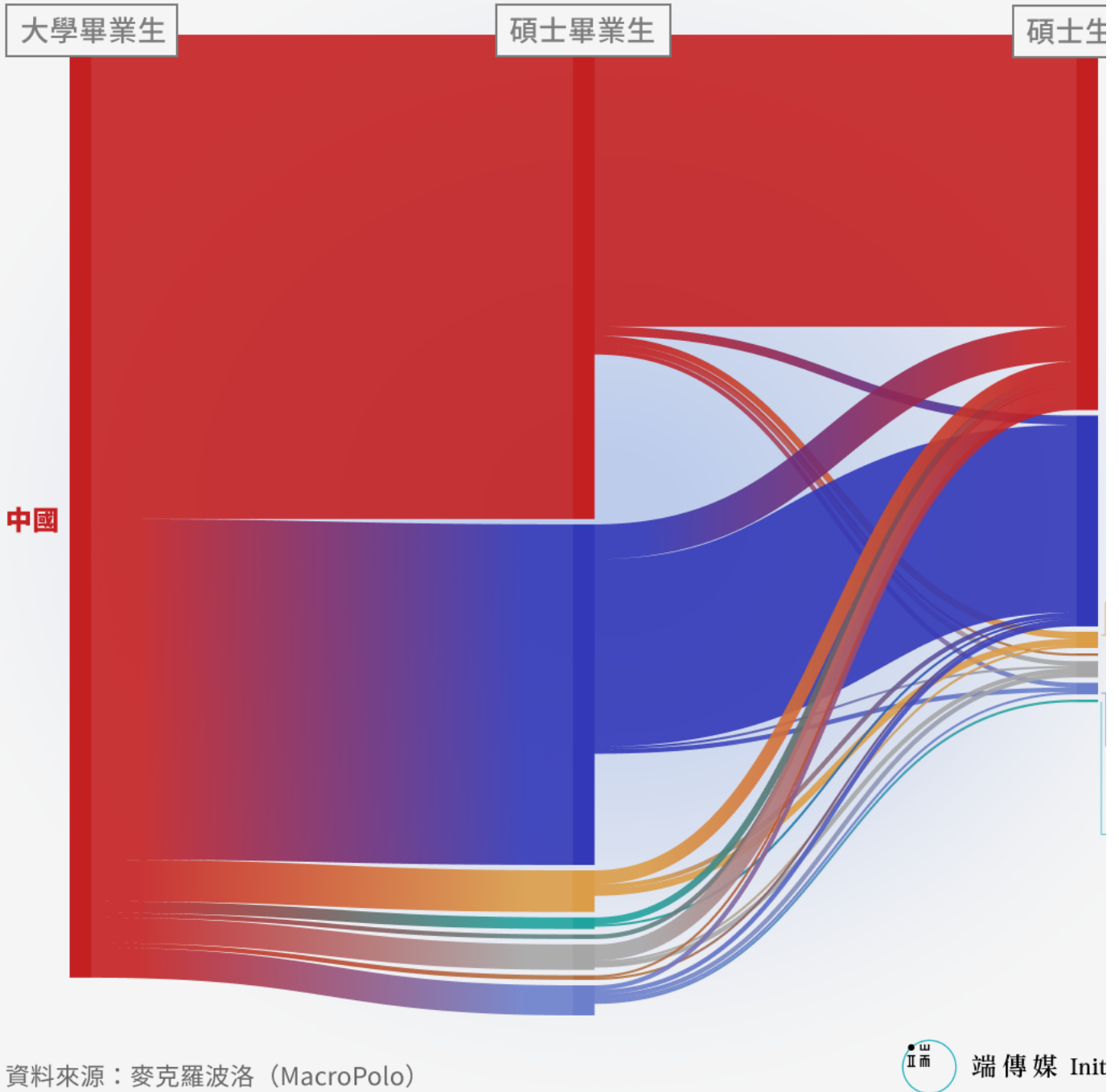
肖文泉认为，芯片禁令只是推高了中国企业的成本。比如，同样一个模型，美国用一个芯片就能跑，中国买不到最快的芯片，但可以用几个稍微降级的芯片替代。

「这增加了算力的价格，但也只是一个价格问题。」她同时指出，随着 AI 领域的技术发展，训练模型需要的算力也可能减少。

相比算力、数据上的落后，更难以克服的是人才短缺问题。咨询公司麦肯锡2023年的报告预测，2030年中国的 AI 人才缺口将达到**400万**。

中国是全球 AI 人才的最大产出国。智库麦克罗波洛 (MacroPolo) 追踪全球 AI 人才流动**发现**，约三分之一的全球顶级 AI 人才在中国完成本科教育，他们中有近一半的人在毕业后赴美、英、澳等国修读硕士。而在美国修读硕士的所有学生（含中国留学生）中，超过80%留在美国，另有6.74%的人去了中国工作。

# 中国 AI 人才外流到哪些国家



肖文泉认为，中国当下的市场环境、就业形势都难以吸引到最尖端的人才。

在薪资水平上，中国落后美国一大截。美国薪资查询网站Salary显示，初级机器学习工程师的平均年底薪约12万美元（约等于86.7万人民币）。而中国招聘网站猎聘的数据显示，2024年一季度生成式 AI 领域招聘算法工程师的平均年薪逾48万人民币。

张路洲在美国一家顶尖私立大学获得计算机博士学位，毕业后留在美国工作。他告诉端传媒，他的中国同学中有80%选择在美国就业。「机会和回报都很好，最好的公司也都在美国。」张路洲的选择里还有另一层考量，他不喜欢中国当下的政治环境。

除了留学生不愿回国，一些在中国知名互联网企业或科研机构工作的人，也正考虑离开。肖文泉频繁收到他们的询问：如何在美国申请硕士、如何办理移民？

肖文泉认为，人才匮乏是限制中国底层模型发展的核心因素。「我在 OpenAI 感受最深刻的是，大语言模型——至少现在的 Transformer 模型（注：ChatGPT就是基于Transformer模型构建的具体应用），本质上还是个大力出奇迹的事，你要聚集很多人一起做才有成功的概率。」她说，OpenAI、Anthropic、Deepmind 等公司，基本垄断了整个美国生成式 AI 行业的顶尖人才。而中国没办法聚集这么多人才，却已陷入「百模大战」。

被分散的不仅仅是人力。复旦大学计算机学院教授邱锡鹏指出：在数据、算力和研发资金都有限的条件下，每家公司依然关起门来各做各的，「做一些很基础的、重复性的事情」。

截至发稿前，中国已有超过4500家 AI 公司，发布逾300个生成式 AI 模型，其中约140个获得官方准许面向大众开放使用，百度、阿里、腾讯、抖音等互联网大公司均名列其中。

不过，一个被媒体广泛报导的事实是：不少模型依赖美国的底层系统，也就是业内说的「套壳」——在已有的开源模型（如脸书母公司Meta的模型Llama 3）的基础上进行微调，产出一个新模型。

在数据、算力、人才乃至底层模型都落后的情况下，多数中国公司将「宝」押在了应用上。中国 AI 公司百川智能创始人王小川的话被广泛流传：当下中国 AI 公司需要思考的是，如何用一个稍弱的模型做出一个好的 AI 原生应用。

这一逻辑，和中国在移动互联网时代的发展逻辑类似——在已有的 iOS、Android 系统内开发不同功能的App。中国公司确实创造出不少用户量惊人的App，在产品体验、市场运营和把握用户需求等方面积累了大量经验。

「漫长的技术发展经验告诉我们，这已经不是中国第一天面对基础科技能力不如国外的棘手问题……暂时的技术领先也并不代表什么。」一篇讨论中国生成式 AI 产业未来的报导写道。

张路洲认为：「如果一条道路确定能够成功，只需要堆人力尝试各种可能性、找到最佳解法，中国公司往往能做得很好。」但生成式 AI 在发展过程中还会经历至少一两次重大的范式转移，将游戏规则全部改写。他觉得，这种重大的技术突破更有可能发生在美国，后者有更宽松的社会文化、更密集的人才，最重要的是——愿意投资目前看不到回报的东西。

不过，在下一个范式转移到来前，中国的科技公司要先抓住这一轮的技术红利期，在市场上占据一席之地。

2023年3月16日，百度首席执行官李彦宏在北京举行的新闻发布会上介绍文心一言的功能。摄：Ng Han Guan/AP/达志影像

## 中国互联网公司的命运「轮回」

目前，中国科技公司尚未收获和信心相匹配的成果。

据数据调研机构 QuestMobile 统计，截至2024年3月，中国生成式 AI 的 App 月活跃用户总数超过7380万。其中有两款 App 的月活跃用户突破千万，分别是抖音旗下的「豆包」（2328万）和百度「文心一言」（1466万）。

与之相较，ChatGPT 截至2024年3月的月活跃用户已超过1.8亿。

业界认为，中国生成式 AI 起步较晚，还要面临监管和内容审查，以文心一言为例，它发表于2023年3月，但直到8月才获政府放行、向大众开放使用——比 ChatGPT 的面世迟了九个月。因此，中国还需要一段时间培养市场，特别是用户的付费意愿。不过，在中国 AI 创业者许德昌看来，各家公司还没赚到钱，却已走上移动互联网时期的「老路」——烧钱、圈地、再考虑变现。「你不免费，竞争对手也会抢先免费，把你的潜在客户拿走。所以大家都先别考虑赚钱了，先赔钱把用户圈到手里。」许德昌说。

面向B端的服务也是卷到不行。

服务B端的思路是，借助生成式 AI 的能力为企业制定方案、升级系统、优化流程等，进而降低成本、提高生产力。比如，华为的盘古大模型3.0面向政务、煤矿、制药、气象等领域，网易的「子曰」面向教育，百度「灵医」面向医疗。

尽管各家公司都在宣扬 AI 业务拉动了收入增长，一篇名为《大模型创业，谁赚到钱了？》的**报导**却戳破了泡沫：「仔细剖析你会发现，业务还是之前的业务，只是换了个名头，加了个帽子而已。」报导采访的一位投资人表示，自己仍在试图搞清楚，哪些公司将把 AI 的前景转化为长期利润。他用冰球比赛来类比：中间的冰球没有被控制住，没有人知道它会去哪里。

许德昌指出，针对医疗、金融等领域的定制化 AI 服务都有相当长的项目周期，目前尚未出现有影响力或赚钱的案例。况且，很多项目都是为了拿一个好看的客户案例，借此吸引行业中的中小企业客户，利润是很薄的。

「To C 的产品还没发展出一个非常成熟的路径，To B 的市场很小。」肖文泉说。

资本也透出犹疑。研究机构CB Insights发布的《2023年人工智能（AI）行业现状报告》显示，2023年，美国AI领域初创公司的投融资数量约为1151笔，总融资达310亿美元，同比增长14%。而中国AI领域投融资数量约为232笔，融资总额约为20亿美元，同比下降70%。

「现在经济环境不是特别好，大家对国内创业的企业持悲观态度。」肖文泉说。

不过，许德昌认为，针对 AI 的投资已经是中国互联网行业持续下行趋势下的一波「逆流」了。「如果没有 AI 这一波，去年和今年应该会更冷。」

中国科技公司希望在 AI 时代复刻移动互联网时期的成功。和上一波浪潮一样，他们依赖西方底层技术、擅于创新。但不同的是，西方的技术和政策不似过去那样开放和友善；出于地缘政治、经济下行等原因，海内外的资本不复当年的热情；在一个技术被「卡脖子」、人口红利殆尽、消费萎靡的中国，每一步都变得更难行。

许德昌举了个例子：推出一个新App，过去获客成本（指为了获取新客户所需的销售及市场营销成本总额）是两块钱人民币一个用户，「现在你五十块钱都买不到一个真实用户」。

唯一没有变的，是防火墙。

中央电视台开始在新闻视频中使用 AI 生成的视频，更推出由生成式 AI 制作的动画片《千秋诗颂》。图：网上图片

迄今，在中国境内推出的生成式 AI 模型绝大多数都需要使用中国大陆手机号码注册帐号，方可使用。与此同时，包括 ChatGPT、Gemini、Copilot、Claude等主流生成式 AI 模型均不对中国用户开放。以ChatGPT为例，和中国一起被其列入禁用国家的还有俄罗斯、伊朗、北韩等。

这一割据局面或早已注定。互联网在1994年进入中国，并在随后30年长成一个繁荣、封闭的世界，新技术可以进入这个世界，却无法打破其封闭性。

深谙这一点的中国科技公司，将鸡蛋分别放在了墙内和墙外两个篮子里。

据「非凡产研」[数据](#)，截止2024年4月，中国出海规模以上（指月访问量超过1万）的 AI 产品数量已达59个。

这一方面是因为海外市场环境更友好，用户付费意愿更高、获得投资的机会更多、也有更多更先进的基础模型——比如中国创业者在国内不能合法使用的 ChatGPT。

另一个原因则是中国政府设置了包括内容审查、备案等一系列准入门槛，大大增加了企业的时间成本。

因此，不少中国公司会选择在国内外发布不同的版本，海外先行一步，试水温、收集反馈，再回来推出国内版。比如初创公司 MiniMax 分别向国内和国外发布两款 AI 社交 App ——星夜和 Talkie 。后者自2023年6月上线以来，累计下载将近1400万次。

地缘政治的阴影笼罩着出海的中国企业。肖文泉指出，很多美国主流基金因为担心政治风险，不愿投资出海的项目，特别是B2B赛道。不少中国公司的研发团队身处中国境内，也增加了美国公司的戒心。

「我见过绝大多数的出海公司都是一方面不愿意拿国内VC（风投）的钱，一方面拿不到美国VC的钱，比较尴尬。」肖文泉说。中国创业者都心知肚明，拿了中国的钱，就不能再拿美国的。

以初创企业 HeyGen 为例，该公司在中国累计获得近千万元人民币的投资。出海后，HeyGen 注销了在中国的公司主体，并于去年11月在美国完成一轮总规模560万美元的融资，估值抬升至7500万美元。HeyGen 制作的泰勒·斯威夫特、特朗普讲普通话的视频曾被广泛传播。

「你很难再看到中国人主动站出来说我在海外赚多少钱、做得多好。」许德昌说，很多海外创业者们「不说自己是中国人、不说自己是中国的项目，尽量避免有中国元素」。

2024年1月12日，北京，一名男子在人行天桥上一边抽烟一边看手机。摄：Florence Lo/Reuters/达志影像

## 尾声

尽管行业非常卷，许德昌依然对中国市场有信心，其中一个原因来自政策利好。

今年两会发布的《2024政府工作报告》首次提出「人工智能+」行动，以政府之力推动 AI 技术在各行业的运用。中国官媒将「人工智能+」与2015年政府工作报告的「互联网+」相提并论，它们都是政府主推的产业升级策略。

「各种行业峰会都在说要拥抱AI，大家就会形成一个我要是不去做，会显得我不太积极、不太作为的逻辑。」许德昌说。

一些行业已经走在了前面。中央电视台开始在新闻视频中使用 AI 生成的视频，更推出由生成式 AI 制作的动画片《千秋诗颂》、由 AI 译制的英文版系列微纪录片《来龙去脉》等。据媒体不完全[统计](#)，上海、长沙、北京等地的广播电视台先后发布、正在筹划的AI作品或有几十部。

这也形成了一个生态闭环：部分中国 AI 公司在饱受严苛监管与中美脱钩困扰的同时，受益于官方的政策扶持和防火墙内庞大的中国市场，并服务于最大的甲方——政府。

一些企业已经尝到「甜头」。3月底传出[消息](#)，Apple正在与百度商讨，在中国市场的苹果设备中使用百度的生成式 AI 技术。因为当 Apple 要为中国区 iPhone 用户提供 AI 功能时，它必须从获得中国政府备案的中国公司里挑选合作对象。此前，三星在中国推出Galaxy S24 系列手机时，也是用文心一言换掉了 Gemini 。

我们正在迈入一个由 AI 赋能的未来，此刻即是确立未来游戏规则的时间窗口，包含监管、价值观、发展策略及运用范围。在中国，这一切的确立都非常高效。

回到最初的问题：如何做一个「合格」的中国生成式 AI ？答案或许是，成为「墙」的一部分。

应受访者要求，许德昌、张路洲为化名。

[#生成式人工智能#chatgpt#中国人工智能监管#AI#内容审查#人工智能](#)

本刊载内容版权为端传媒或相关单位所有，未经[端传媒编辑部](#)授权，请勿转载或复制，否则即为侵权。