

市场简报：深度学习框架到底是什么？ 中国AI企业为何纷纷开源自研框架？

Briefing Report: What exactly is a deep learning framework? Why do Chinese AI companies open up their own self-developed frameworks?

市場速報：ディープラーニングフレームワークとは正確には何ですか？ なぜ中国のAI企業は独自に開発したフレームワークを開くのですか？

报告标签：人工智能、深度学习、开源、框架

主笔人：谢子博

报告提供的任何内容（包括但不限于数据、文字、图表、图像等）均系头豹研究院独有的高度机密性文件（在报告中另行标明出处者除外）。未经头豹研究院事先书面许可，任何人不得以任何方式擅自复制、再造、传播、出版、引用、改编、汇编本报告内容，若有违反上述约定的行为发生，头豹研究院保留采取法律措施、追究相关人员责任的权利。头豹研究院开展的所有商业活动均使用“头豹研究院”或“头豹”的商号、商标，头豹研究院无任何前述名称之外的其他分支机构，也未授权或聘用其他任何第三方代表头豹研究院开展商业活动。

Q1: 如何实现深度学习算法的开发与应用？深度学习框架对人工智能应用的开发及行业的发展所起到的作用是什么？

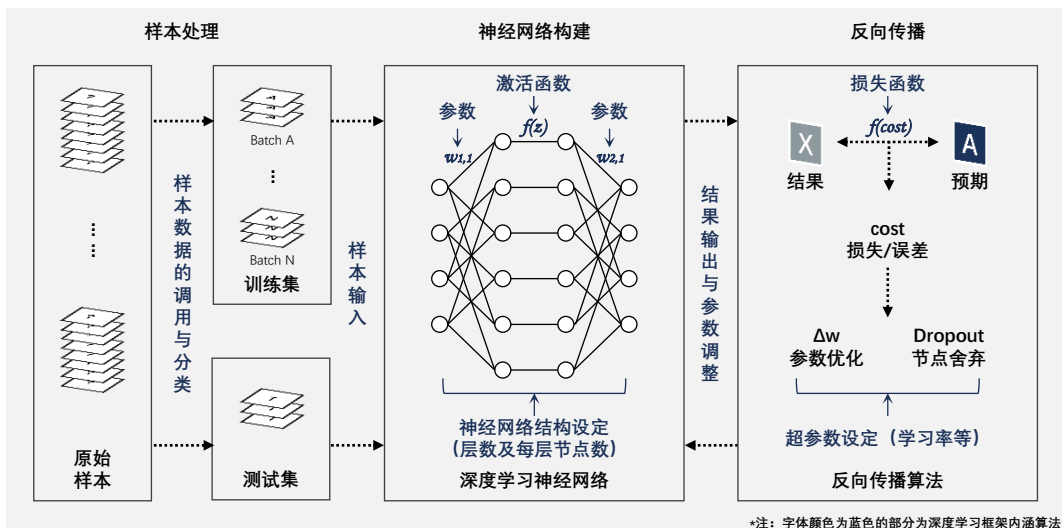
■ 深度学习框架打破技术门槛，打通技术交流渠道，促进业界协同发展

深度学习是以人工神经网络为基础、融合数学、计算机科学、神经科学等多种学科、对数据及信息进行学习的机器学习算法。深度学习底层算法的开发对算法开发人员的要求较高，需精通多个前沿科学领域理论并具备相对应的应用能力。由于较高的学术及技术门槛，在深度学习技术发展早期缺乏底层开发工具的情况下，深度学习技术发展速度较慢，相关应用的落地难以实现。

深度学习框架的出现打破了基于深度学习技术的人工智能应用开发的高技术门槛。深度学习框架是一系列经过高级计算机语言封装的深度学习底层算法，内含可实现样本数据的调用、分类及输入、神经网络架构建设、激活函数及损失函数等函数设置、超参数设定及调整等功能的完整函数工具。通过对深度学习框架的引用，人工智能应用开发者无需熟知底层算法开发逻辑即可完成应用开发，人工智能应用开发的技术门槛大幅降低。

深度学习框架的出现打通了业界技术交流与知识分享的渠道。采用同一深度学习框架的算法程序具备较高的易读性及兼容性，使算法程序的移植及调用难度大幅降低。算法开发人员可通过学习已开源的算法区块，并对该算法进行调用或进一步研发，在已有知识架构体系的基础上实现人工智能算法的创新。

图表1：深度学习框架概述



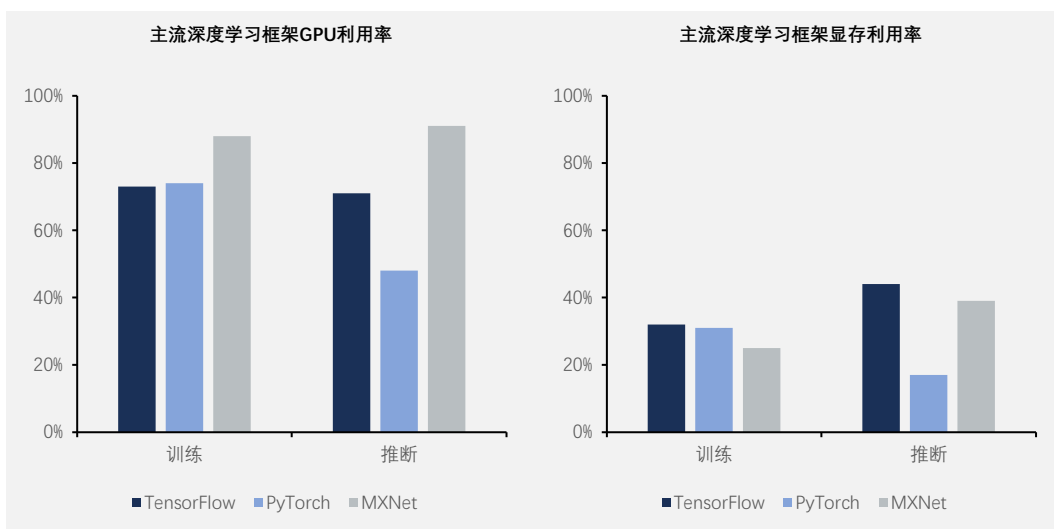
Q2：在目前市场上已存在可开源利用的、由海外企业开发的成熟框架的背景下，中国人工智能企业为何纷纷布局深度学习框架的自主研发及推广？

■ 现存深度学习框架在效率上的缺陷及日趋复杂的国际形势驱动中国人工智能企业布局深度学习框架自主研发

中国人工智能企业布局深度学习框架的自主研发及推广的主要原因有二。一是目前业界主流的深度学习框架在运算效率上尚有缺陷，主要体现在算法训练及推断阶段对GPU、CPU等硬件的利用率较低。如何优化深度学习框架编写逻辑及计算机语言编译过程，从而提升深度学习框架对硬件的利用率成为全球人工智能学术界及产业界亟需解决的问题，也成为企业抢占人工智能行业高地的机会。具备更高效率的深度学习框架将成为替代现存主流深度学习框架、推动企业人工智能生态构建的有力支持。

第二个原因则归咎于目前全球日益复杂的贸易环境。主流的深度学习框架虽已由全球各大人工智能研发主体开源发布，其知识产权仍由各研发主体持有。受国际形势影响，已开源的深度学习框架或将面临商用受限等风险。由于TensorFlow、PyTorch等由海外主体研发并开源的深度学习框架已占据业界主流地位，框架商用受限将对业界生态、行业后续发展及人工智能企业存续产生显著的负面影响。国际关系的日趋复杂迫使各国政府及人工智能研发主体加强包括深度学习框架等核心技术的自主研发，从而掌握人工智能领域发展的主动权。

图表2：主流深度学习框架对GPU及显存的利用率，基于计算机视觉算法



来源：头豹研究院编辑整理

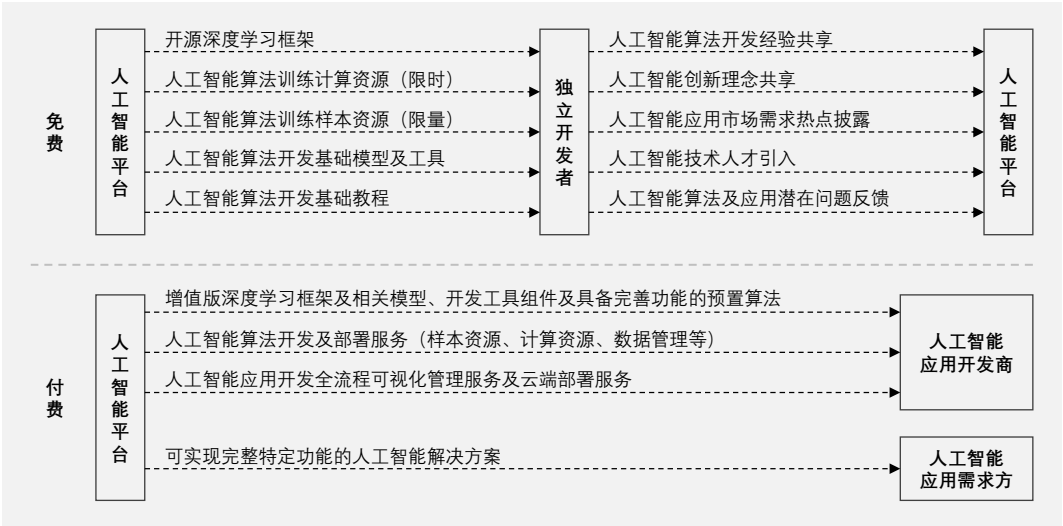
Q3：在知识产权愈发受到重视的今天，全球人工智能学术界及产业界为何纷纷开源旗下自研的深度学习框架？

■ 开源是促进技术创新、业务增长及行业发展的关键驱动力

与程序及软件的开源类似，深度学习框架的开源指框架以开源许可证形式发行，任何主体均有权进行查看、在其源代码基础上根据需求进行修改（生成新的代码）并加以应用。围绕深度学习框架的开源，人工智能企业基于其强大的云服务实力及人工智能技术研发实力，针对人工智能应用开发流程的简化及开发环境的优化为业界提供一系列人工智能开发工具及服务，为人工智能业界的技术创新及协同发展提供有力支持。同时，深度学习框架的开源与平台的建立为人工智能企业本身的技术创新、产品优化、应用拓展、人才引进提供有力支持。深度学习框架的开源是建立成熟AI生态的核心引擎。

此外，深度学习框架的开源利好人工智能平台运营商付费业务的拓宽及增长。在以开源深度学习框架为核心的免费服务基础上，人工智能平台运营商保留面向人工智能应用开发商及需求方的付费服务。付费服务主要围绕算法、算力、数据与应用部署服务为业界提供人工智能全流程开发支持，是人工智能平台运营商的主要收入来源。深度学习框架的开源助力人工智能技术研发及应用落地的同时，推动人工智能生态的构建及应用领域的拓宽，释放各应用领域智能化转型需求，为人工智能平台运营商付费业务的增长提供有力支持。

图表3：人工智能开发平台商业模式

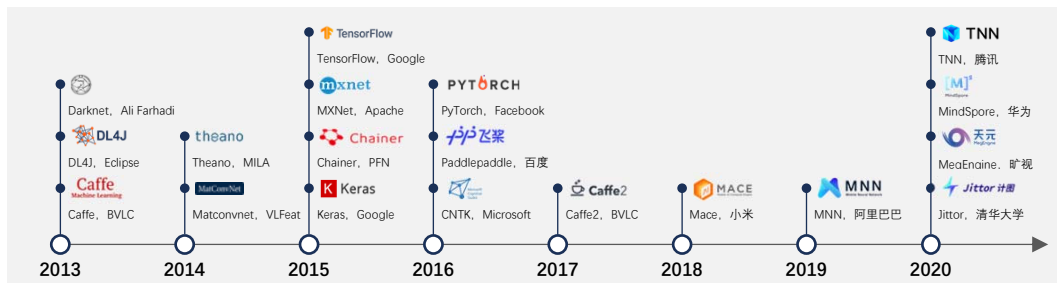


来源：头豹研究院编辑整理



Q4：已开源的深度学习框架有哪些？主流深度学习框架由哪些公司开发？中国人工智能企业自研的开源框架与主流框架的差距在哪里？

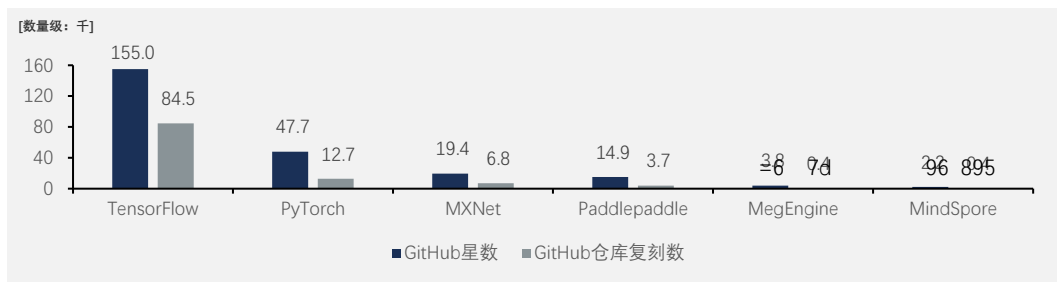
图表4：深度学习框架开源简史



■ 深度学习框架开源由来已久，TensorFlow及PyTorch占据业界主导地位

自2013年始，全球人工智能学术界及产业界各研发主体陆续开源旗下自主研发深度学习框架，并以框架为主体搭建人工智能开放开发平台，推动人工智能产业生态的建立。其中，由Google团队开发的TensorFlow及由Facebook团队开发的PyTorch两款深度学习框架占据业界主导地位。

图表5：深度学习框架对比，截止至2021年4月19日



■ 中国企业自研深度学习框架短期内实现国产替代难度较大

凭借较高的兼容性及应用难度，TensorFlow、PyTorch两款深度学习框架成为业界开发者的主流选择。由百度、小米、阿里巴巴、腾讯、华为、旷视科技等中国人工智能企业自主开发并开源的深度学习框架起步较晚，在社区繁荣度、开发者数量、说明文件丰富度等层面均不及TensorFlow与PyTorch，短期内难以实现对海外框架的全面替代。

方法论

- ◆ 头豹研究院布局中国市场，深入研究10大行业，54个垂直行业的市场变化，已经积累了近50万行业研究样本，完成近10,000多个独立的研究咨询项目。
- ◆ 研究院依托中国活跃的经济环境，从人工智能、底层框架、开源生态等领域着手，研究内容覆盖整个行业的发展周期，伴随着行业中企业的创立，发展，扩张，到企业走向上市及上市后的成熟期，研究院的各行业研究员探索和评估行业中多变的产业模式，企业的商业模式和运营模式，以专业的视野解读行业的沿革。
- ◆ 研究院融合传统与新型的研究方法，采用自主研发的算法，结合行业交叉的大数据，以多元化的调研方法，挖掘定量数据背后的逻辑，分析定性内容背后的观点，客观和真实地阐述行业的现状，前瞻性地预测行业未来的发展趋势，在研究院的每一份研究报告中，完整地呈现行业的过去，现在和未来。
- ◆ 研究院密切关注行业发展最新动向，报告内容及数据会随着行业发展、技术革新、竞争格局变化、政策法规颁布、市场调研深入，保持不断更新与优化。
- ◆ 研究院秉承匠心研究，砥砺前行的宗旨，从战略的角度分析行业，从执行的层面阅读行业，为每一个行业的报告阅读者提供值得品鉴的研究报告。

法律声明

- ◆ 本报告著作权归头豹所有，未经书面许可，任何机构或个人不得以任何形式翻版、复刻、发表或引用。若征得头豹同意进行引用、刊发的，需在允许的范围内使用，并注明出处为“头豹研究院”，且不得对本报告进行任何有悖原意的引用、删节或修改。
- ◆ 本报告分析师具有专业研究能力，保证报告数据均来自合法合规渠道，观点产出及数据分析基于分析师对行业的客观理解，本报告不受任何第三方授意或影响。
- ◆ 本报告所涉及的观点或信息仅供参考，不构成任何证券或基金投资建议。本报告仅在相关法律许可的情况下发放，并仅为提供信息而发放，概不构成任何广告或证券研究报告。在法律许可的情况下，头豹可能会为报告中提及的企业提供或争取提供投融资或咨询等相关服务。
- ◆ 本报告的部分信息来源于公开资料，头豹对该等信息的准确性、完整性或可靠性不做任何保证。本报告所载的资料、意见及推测仅反映头豹于发布本报告当日的判断，过往报告中的描述不应作为日后的表现依据。在不同时期，头豹可发出与本报告所载资料、意见及推测不一致的报告或文章。头豹均不保证本报告所含信息保持在最新状态。同时，头豹对本报告所含信息可在不发出通知的情形下做出修改，读者应当自行关注相应的更新或修改。任何机构或个人应对其利用本报告的数据、分析、研究、部分或者全部内容所进行的一切活动负责并承担该等活动所导致的任何损失或伤害。