

中国发展高层论坛 2024 年年会 “持续发展的中国”

人工智能发展与治理专题研讨会 (主题发言二)

由国务院发展研究中心主办、中国发展研究基金会承办的中国发展高层论坛2024年年会，于2024年3月24-25日在北京钓鱼台国宾馆举行。3月24日下午，举行了“人工智能发展与治理专题研讨会”，其中，“主题发言二”环节的发言嘉宾为北京智源研究院创始人、创始理事长张宏江。本场会议由国务院发展研究中心副主任张顺喜主持。

今天大热的多模态大模型，未来发展会是什么样?张宏江认为，其未来发展方向一定不光是做视频生成、视频剪辑、拍电影或是生成电视剧。从技术角度观察，我们可以用它来做机器的大脑、识别外围的世界、武装未来的自动驾驶，从而使我们今天的信息系统、模型系统变成未来的行动系统。

在张宏江看来，行动系统，尤其是多模态大模型最让人振奋的是它可以给机器人一个大脑。比如，给机器人指令从桌上的一堆玩具中抓出一个已经灭绝的动物，它能够经过推理和识别过程，从所认识的老虎、狮子、鸟等若干种动物中，成功地抓出已经灭绝的动物恐龙。这是过去的机器人不能做到的。过去你告诉机器人抓什么东西它能抓得到，但是给它一个抽象的概念，它是无法完成的。同样，告诉现在的机器人说渴了，机器人就会从一堆物体中抓出一个有水的瓶子。这两个演示都说明，机器人有了多模态大模型之后，它们不再是之前简

单地听你给它的指令，而是能够思考你指令中的含义，这是我们看到未来自主机器人的雏形。多模态大模型今天已经能够让我们如此震撼。

这背后的一切，实际上是过去70年人工智能持续发展的结果。人工智能经过了三次发展浪潮。在第三次浪潮里又经过过去10年深度学习的浪潮。大模型的出现正好是2020年第三次浪潮处于低谷时，而ChatGPT3.0的出现是一个转折点，它带来了Sora和一系列中国的大模型。在过去十几个个月内，我们看到从语言模型、多模态模型到视觉模型，再到未来大模型的快速发展，这背后的原理是什么？一个很重要的因素就是我们今天做人工智能不再是把它当作一个算法，而是把它当作一个系统来做。今天的模型不光是规模大，而且具备通用性。这背后的驱动力是“规模增长定律”。正是这种规模效应，使它能够解决一个又一个过去我们无法解决的问题。比如语言模型，当它只有几十亿数据时，只能解决原来自然语言处理中的一些问题，但当它超过5千亿数据时，基本上自然语言中所有的问题都迎刃而解，这就是规模的能力。

大模型背后的核心到底是什么？张宏江认为是一个新的操作系统。传统的PC时代，是经过CPU的计算产生输出；今天大模型计算的核心已经不再是CPU，而是GPU，因此说它是一个新的操作系统。今天所有互联网平台公司都在努力地做大模型，其根本的一点是，如果没有大模型，未来他们就不再是一家平台公司。观察过去几年的发展过程，尤其是过去18个月的发展，我们可以总结出新的摩尔定律，即模型的能力每一到两年提升一代，训练的成本每18个月变成之前的1/4，模型的推理成本每4个月会变成前面的1/10。这个新摩尔定律将带来大模型的快速普及、快速发展和快速应用。另一个推动力，就是英伟达的股票在过去12个月的突飞猛涨，该公司已经成为世界前三名最有价值的公司。整个大模型产业链现在都在快速地发展、快速地成长起来。大模型会赋能我们的软件工具，赋能我们的生活，赋能我们的工

作。

当然，未来通用人工智能的发展会产生一些风险，这个风险甚至可能导致人类灭绝。杰弗里·辛顿从去年3月份开始，非常深刻地回顾了过去人工智能和神经网络的发展，其结论是，数字计算本身的效率事实上已经超过了生物计算的效率，因为数字系统可以彼此互相学习，可以把之前积累的东西传到下一个系统里去。这意味着，通用人工智能很快就会在各个方面超越人类智能。

以前人类的学习能力基本上是一条平缓的直线，但是机器的学习能力却是一条指数曲线。随着时间的推移，当这两条线相交之后，人类的学习能力已经赶不上机器的学习能力了。也就是说，机器消灭的工作，依然由新的机器所替代，而不是由人替代。那么，人的工作是不是会彻底消失？大部分的人会不会彻底失业？事实上，失业还不是最可怕的。最可怕的是，当系统本身如果能够有能力重新自我改善，那它所带来的问题，对人类来说就是会产生一些深层次风险。如果想避免人工智能或将导致全球性的灾难后果，那么从今天开始我们就必须采取果断行动，而且必须加强国际合作，无论是技术上、立法上，还是治理方面。

张宏江说，今天我们看到人工智能进入了一个新的发展阶段，以大模型为代表的阶段，代表了第四次技术革命（前面的三次是农业革命、工业革命、信息革命），将带来大量的效率提高，为我们的生活提供非常多的方便，会创造巨大的价值和又一个新的产业。但与此同时，我们也要看到人工智能可能导致的全球性的灾难性后果，这种后果可能是人类灭绝性的危险。为了避免这样的危险发生，我们需要划出一些红线，改进治理机制；同时还要研发更多的安全性技术，控制人工智能不去跨越这些红线。要想做到这一点，最重要的一条就是我们要坚持和加强国际科学界和政策界在安全方面的合作，只有这样，我们才能避免这场灾难的发生。

(中国发展出版社 撰稿：刘长杰；审稿：杨良敏
中国发展研究基金会 审校：丁孟宇)

--背景信息--

中国发展高层论坛是由国务院发展研究中心主办、中国发展研究基金会承办的国家级论坛。论坛秉承“与世界对话，谋共同发展”的宗旨，着眼于对外释放我国对外开放信号，自2000年创办以来，已发展成为服务中央决策和国家对外经济交往、促进中外企业交流合作的重要平台。

--媒体联络--

中国发展研究基金会

郭丝露 18666028168/64255855-8014

史雅帆 13810361966/64255855-8223

夏天 18801375838/64255855-8086

施婉菁 18801090391/64255855-8090

邱凯娴 18301078627/64255855-8103



中国发展研究基金会
China Development Research
Foundation

中国北京东城区安定门外大街138号
皇城国际中心A座15层

邮编: 100011
电话: 86-10-64255855
传真: 86-10-64255855-8100
网址: www.cdrf.org.cn
电邮: cdrf@cdrf.org.cn

Floor15, Tower a, Imperial International Center, No.138
Andingmen Wai Avenue, Dongcheng District, Beijing, 100011, China

Tel: 86-10-64255855
Fax: 86-10-64255855-8100
Website: www.cdrf.org.cn
E-mail: cdrf@cdrf.org.cn