

# 什么是第三代人工智能

自 1956 年达特茅斯会议以来，人工智能经历了 60 多年的风风雨雨。在《中国科学》70 周年纪念专刊上，清华大学张钹院士等撰文《迈向第三代人工智能》，提出了第三代人工智能的概念[1]。

按张钹老师的观点，人工智能技术的发展总体可以分为两个阶段：基于知识驱动的方法和基于数据的方法，分别称为第一代人工智能和第二代人工智能。

在第一代人工智能中，人类的知识具有核心地位。人们将定理、概念、经验等知识形式化为符号，并设计合理的推理方法来获取问题的答案。这些知识由人类总结出来，具有明确的意义，且具有可组合性，因此“和人类理性智能一样具有可解释性”。**这一方法的局限在于对知识进行形式化比较困难，很多时候“只能就事论事地解决特定问题”**。如何引入常识，如何处理不确定性，如何对原始数据进行结构化，都是具有挑战性的问题。

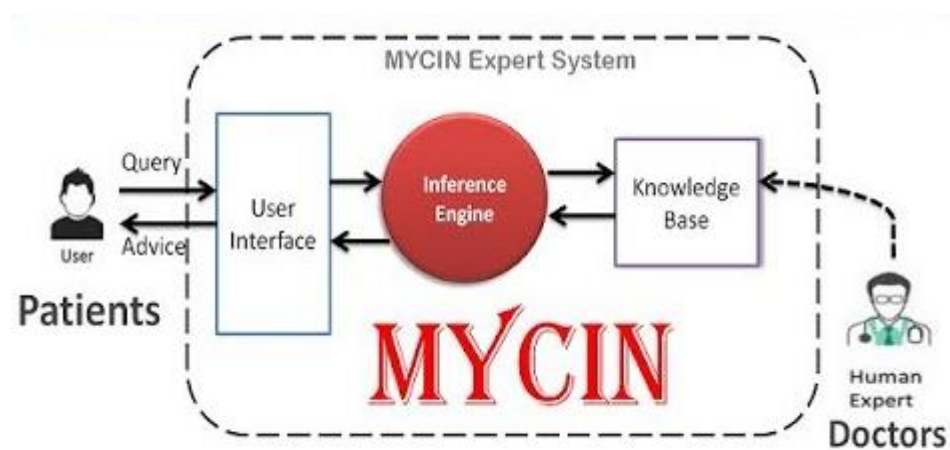


图 1：第一代人工智能举例：MYCIN 诊疗专家系统[3]

在第二代人工智能中，数据具有核心地位，人们不再向机器灌输知识，而是设计算法让机器从数据中自动学习知识。神经网络模型是典型代表，这一模型具有强大的学习能力，可以从原始数据中抽取知识并存储在网络参数中。近年来，随着数据的不断积累和计算机性能的提高，这一方法取得了巨大成功，在某些方面（如图像识别）甚至超过了人类水平。然而，这种纯数据驱动的方法具有明显的“黑箱”特性，人们很难理解学习过程和学习结果，系统的行为也很难控制。

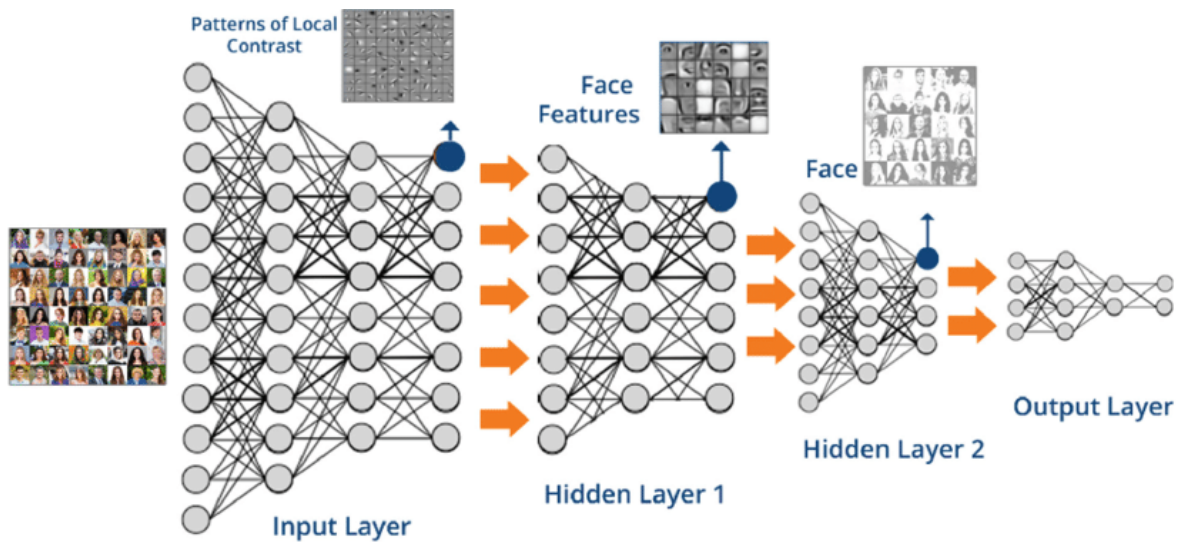


图 2：第二代人工智能举例：基于深度学习的人脸识别系统[2]

第三代人工智能的基本思路是融合知识驱动和数据驱动的优点，“建立鲁棒与可解释的 AI 理论与方法，发展安全、可信、可靠与可扩展的 AI 技术”。换句话说，新一代人工智能会全面利用知识和数据，让 AI 系统不仅可以为我们服务，而且可以明明白白地为我们服务，更强大，更安全，更可靠。

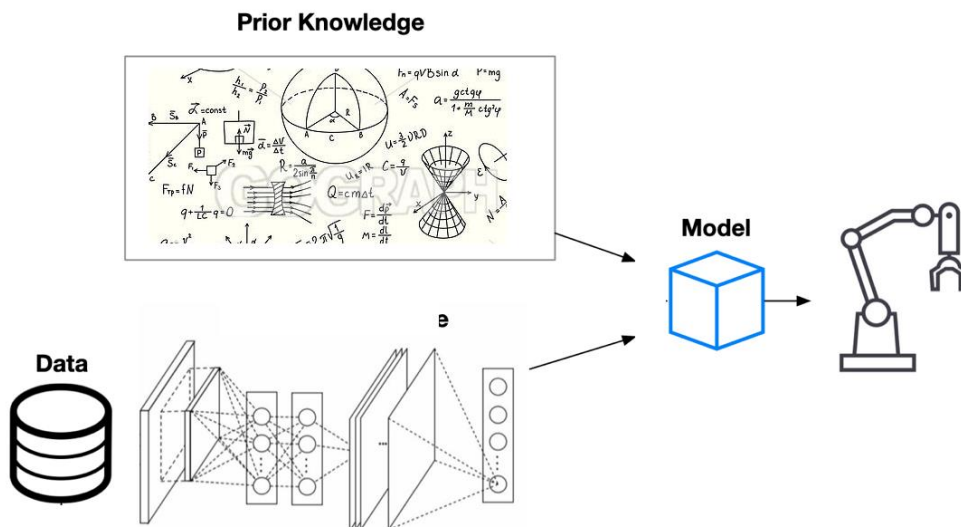


图 3：第三代人工智能：知识与数据的结合

[1] 张钹, 朱军, 苏航. 迈向第三代人工智能. 中国科学: 信息科学, 2020, 50: 1281-1302, doi:10.1360/SSI-2020-0204 Zhang B, Zhu J, Su H. Toward the third generation of artificial

intelligence (in Chinese). *Sci Sin Inform*, 2020, 50:1281-1302, .

[2] Grigsby, Scott. (2018). *Artificial Intelligence for Advanced Human-Machine Symbiosis*.  
10.1007/978-3-319-91470-1\_22.

[3] *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Edited by Bruce G. Buchanan and Edward H. Shortliffe, 1984.