

第二章 AI大模型的技术内核

深入探究信息密度提升的实现路径：从架构突破、智能涌现、对齐驯服到最终的开源创新爆发。

人工智能发展简史

○ 符号主义 (1950s-80s)

核心思想：“智能即逻辑”，构建“专家系统”。

○ 统计机器学习 (1980s-2000s)

核心思想：“智能即归纳”，从数据中学习规律。

○ 深度学习 (2010-2018)

核心思想：“智能即特征层次”，自动学习复杂特征。

○ 大模型时代 (2018-至今)

核心思想：“智能即规模”，催生“涌现”能力。

驱动力的范式转变

摩尔定律



由硬件物理极限驱动，关注晶体管密度，赋能“计算”。

规模定律



由模型、数据、算力三大要素驱动，赋能“智能”。

信息密度提升的四步实现路径

1. 架构革命：挣脱先验束缚

Transformer架构通过其核心的“自注意力机制”，放弃了传统CNN/RNN架构中对数据局部性或时序性的“先验假设”。它用巨大的计算量进行全局计算，换来了前所未有的信息表征密度和灵活性，为承载高密度信息创造了可能。

信息捕捉能力对比

CNN / RNN



受限于局部“感受野”或线性时序，难以捕捉长距离依赖。

Transformer



可直接计算序列中任意两点关系，实现全局信息捕捉。

智能涌现的“相变”过程

临界点

“涌现”发生

模型能力

模型规模 →

2. 规模的暴力美学：智能“涌现”

“规模定律”(Scaling Law)揭示了“大力出奇迹”的可预测性。当模型的参数量、数据量和计算量同时跨越某个临界点后，会突然表现出未被直接训练过的新能力(如思维链)，这被称为“涌现”——是模型内部信息密度达到阈值后发生的“相变”。

3. 智能的驯服：从原始势能到价值对齐

一个强大的模型分两步诞生：首先通过“预训练”阶段，在海量数据中学习，最大化地吸收人类知识，构建一个强大的“基础模型”；然后通过“对齐”阶段(包含SFT和RLHF)，将其巨大的信息势能引导到符合人类价值观的方向上，使其变得有用、无害。

模型诞生的两个阶段



1. 预训练

吸收海量知识，构建原始智能



2. 对齐

注入人类价值观，使其可控有用

开源模型的衍生创新

开源基础模型



法律



医疗



金融



艺术

4. 生态的“寒武纪”：开源创新爆发

以Llama、DeepSeek等为代表的高性能开源模型，极大降低了AI技术的准入门槛，导致模型数量和种类如“寒武纪生命大爆发”般增长。一个强大的开源基础模型发布后，会迅速衍生出成百上千个面向垂直行业的专用模型，行业瓶颈从“模型供给”转向了“场景需求”。