

龙虾：第三次AI平民化

自主Agent时代到来

肖睿

20260322@北大光华管理学院

模型：智能的本质

生命智能

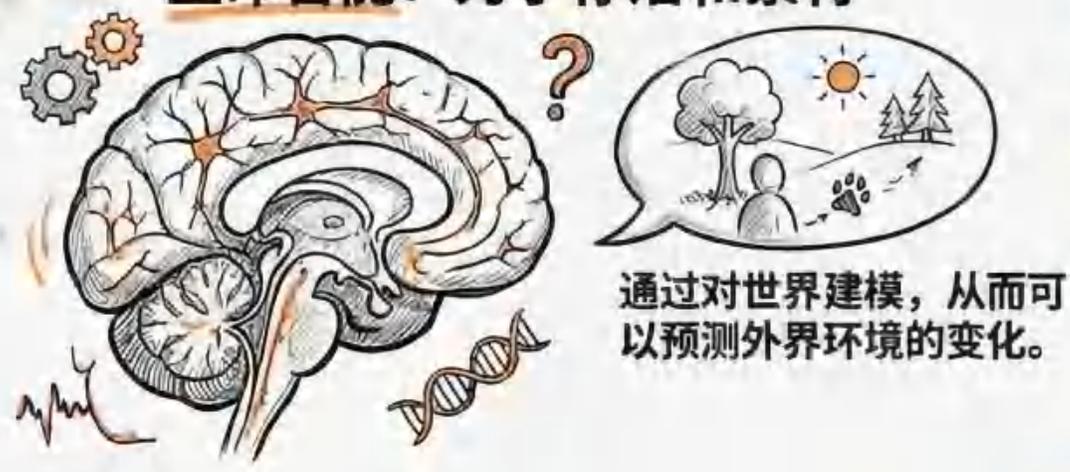
- 生命的目标是存活和繁衍，为了适应外界环境，发展出生命智能。
- 生命智能通过对世界建模，从而可以预测外界环境的变化。

人工智能

- 人工智能对世界建模，从而可以预测下一个数据或动作。
- 现代人工智能既不使用逻辑模型，也不使用数学模型，而是使用数据模型。
- 从统计学习（机器学习），到神经网络（深度学习），最后到大模型（Transformer）

人工智能的本质：从生命进化到数据预测

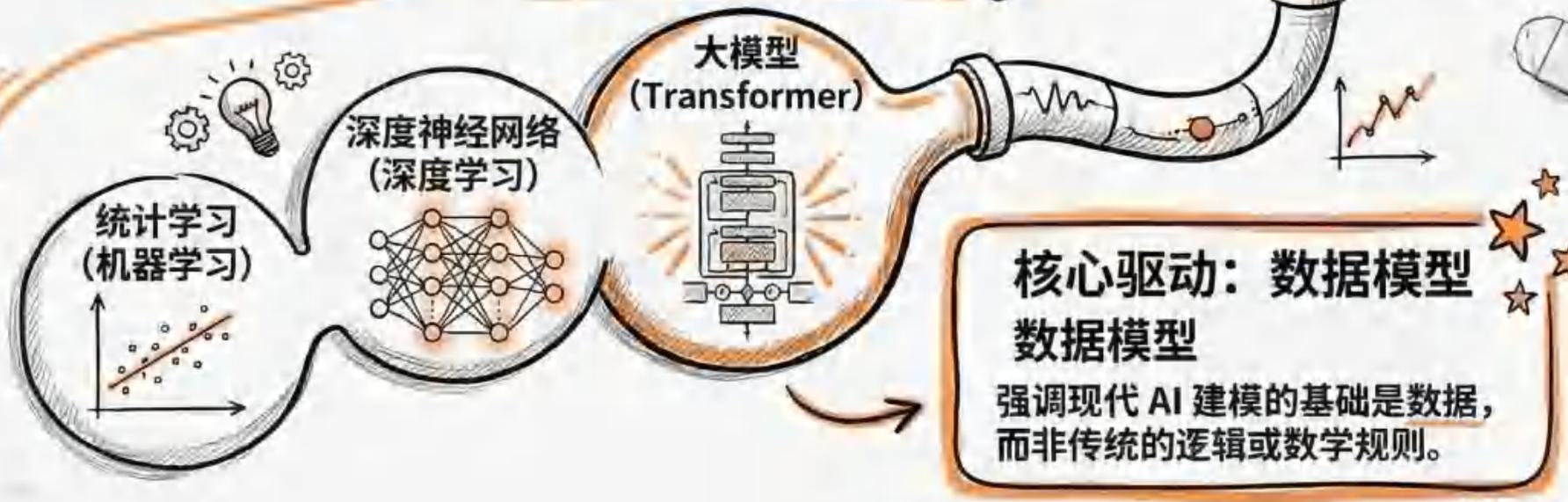
生命智能：为了存活和繁衍



人工智能：预测下一个数据或动作



现代 AI 的技术演进



端到端的表示学习，通过NTP来实现智能

模型2017: Transformer

• 智能如何可能

1. **Token**: 万物皆token
 - 可以计算的语义
2. **Attention**: 智能的来源
 - 最基本的智能算子
3. **GPT**: 智能的公式
 - 用神经网络表示的数学公式逼近世界的真相

• 智能的根基

1. **DATA**: 现代人工智能是数据智能
 - 数据表达世界
 - 模型是数据的有损压缩和泛化
2. **GPU**: 计算GPT公式的机器
 - 批量简单运算的利器

• 智能的使用

1. **Prompt**: 模型的生活环境
 - 人类PUA人工智能的token序列
 - 越聪明的模型需要越简单的Prompt
2. **Agent**: 能动的模型
 - “人类”使用人工智能的抓手
 - 感知、规划、执行、反思、记忆

模型 2017: Transformer 的智能版图

解释 Transformer 模型的核心构成要素及其实现智能的逻辑路径。

智能的核心逻辑 (The Logic of Intelligence)



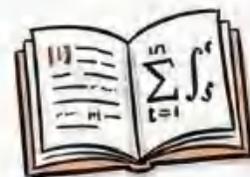
Token: 语义的可计算化

万物皆 Token, 是让语义变得可以被计算的基础单元。



Attention: 最基本的智能算子

它是智能的来源, 构成了模型理解信息的核心机制。



GPT: 逼近真相的智能公式

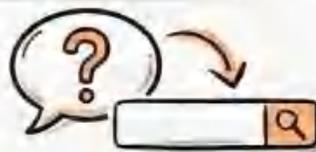
利用神经网络表示的数学公式, 去无限逼近这个世界的真实规律。

运行资源与交互抓手 (Resources & Interaction)



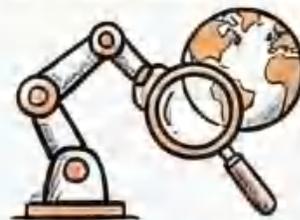
Data & GPU: 世界的表达与运行平台

数据表达世界 (有损压缩), GPU 则是进行批量简单运算的利器。



Prompt: 模型的生活环境

它是人类引导人工智能的 Token 序列, 决定了模型的输出背景。



Agent: 使用 AI 的抓手

作为人类使用人工智能的具象化工具和核心连接点。

现代人工智能已经经历了三次平民化，也走到了第三阶段

模型的未来：AGI的5级路线图

Chat: 对话者

L1

(2020开始-2023出圈)

ChatGPT

- 内容理解和生成
- 知识压缩和泛化
- 指令遵循和性格
- 多轮对话和记忆

Reasoning: 推理者

L2

(2024开始-2025出圈)

DeepSeek

- 深度思考和推理
- 代码生成和反馈
- 知识生产和逻辑
- 复杂问题的解决

Agent: 行动者

L3

(2023开始-2026出圈)

OpenClaw

- 工作流workflow定义
- 工具使用和计算机使用
- 自主Agent: PDCA循环
- 具身智能和物理世界

Innovator: 发明者

L4

(2025开始)

AI4Science?

- 科研和技术创新

Orgnizationner: 组织者

L5

人是目的，不是手段

! 现代人工智能的第三次平民化进程 !

本文概括了现代人工智能平民化的五个阶段：从以 ChatGPT 为代表的内容对话，到以 DeepSeek 为代表的深度推理，再到自主智能体 Agent 的兴起，最终实现科研创新与组织关系的重构。

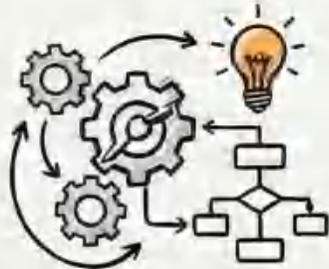
核心智能：从对话到推理 !



Chat (以 ChatGPT 为代表)

- 专注于内容理解与生成、知识压缩与泛化、指令遵循及多轮对话记忆。

Reasoning (以 DeepSeek 为代表)



- 具备深度思考、代码生成反馈、知识生产逻辑及解决复杂问题的能力。

智能形态：智能体与组织重构



Agent (如 OpenClaw, ClaudeCode)

- 实现 workflow 自动化、工具与计算机使用、自主 OPDCA 循环及具身智能。



Innovator & Organizationer  AI4Science

- 驱动 AI4Science 科研创新，并重塑组织观，强调“人是目的，不是手段”。



Agent的进化

非自主Agent

- 扣子Coze1.0
- Dify、n8n、LangChain
- 其他无框架的RAG系统

自主Agent

- 2023年：AutoGPT
- 2024-2025年：DeepResearch、Manus、GenSpark、Lovable
- 2025-2026年：ClaudeCode、**OpenClaw (龙虾)**、Perplexity PC

6-12个月的Agent迭代周期：自主Agent的最新范式就是OpenClaw

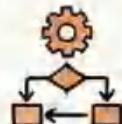
百虾大战：openclaw原生/变体/云端/本地/设备端侧

Agent 的进化图谱: 从辅助工具到自主智能

核心公式: Agent = 感知 + 规划 + 执行 + 反思



感知



规划



执行



反思

该公式定义了 Agent 作为一个完整智能体所必须具备的四个核心闭环环节。

第一阶段: 非自主 Agent

扣子 Coze 1.0

Dify

n8n

LangChain

无框架的 RAG 系统

第二阶段: 自主 Agent (2023-2025)

2023年



AutoGPT

2024-2025

Manus

GenSpark

DeepResearch

Lovable

第三阶段: 自主 Agent (2026 展望)

ClaudeCode

OpenClaw

Perplexity PC

关于OpenClaw (龙虾)

快速出圈：八卦一下

- 出圈 = 惊讶度 X (实用度+便捷度+使用成本) X 运气
 - ◆ 我看到了5次：ChatGPT、DeepSeek、NanoBanana、Seedance、OpenClaw
- 60天github星标超过linux，位居第一
- 无心插柳：Peter40+项目的最后一个，warelay (whatsapp relay)
- 我心明月：Clawdbot致敬ClaudeCode的吉祥物，三次改名

功德无限：自主Agent平权

- 灵活：开源定制，完全掌控，部署灵活 (本地/云端/沙箱)
- 友好：使用方便，记忆强化，主动心跳
- 进化：技能工具，与模型共成长

极客项目：创新与风险

- 原罪：界面难用、个人场景、无商业化考虑
- 快速迭代，快速产品化
 - ◆ 权利让渡，权限管理体系
 - ◆ Skill和工具的安全认证
 - ◆ 虾群管理：1人多虾；1虾多人；多人多虾
 - ◆ log输出 (thinking)：内部活动可见/可信
 - ◆ Agent时代的安卓？龙虾经济、龙虾生态

OpenClaw: 从周末项目到自主 Agent 平权

起源与爆发的公式



无心插柳: 41 个周末项目之一

起始于 Warelly (Whatsapp Relay) 项目的探索。

60 天 GitHub 星标超越 Linux

凭借极高的惊讶度、实用度与便捷度迅速出圈。



出圈公式

惊讶度 × (实用度 + 便捷度) = 成功

这一公式定义了 OpenClaw 及 DeepSeek 等项目的成功逻辑。

OpenClaw & DeepSeek

核心价值与进化挑战

自主 Agent 平权: 开源、本地、进化

支持开源定制与本地部署, 实现技能与模型的共同成长。

自主 Agent 平权



Token



心力

投入与产出的闭环

投入 Token 和心力, 换取技能、工具、记忆与成果。



技能



记忆



工具



成果

进化中的风险



个人极客项目



权限管理弱

目前仍是个人极客项目, 存在权限管理弱等商业化难题。

养虾的自我修养

一个事实

1. 同一个起跑线：

- 没专家
- 变化快
- 平民化
- 门槛低
- 期望高

两个心态

1. 养龙虾就是养孩子：

- 要投入时间、费用、心力、数据/隐私
 - ✓ 时间：200+小时的多样性问答
 - ✓ 费用：Coding场景，龙虾比CC多3倍token
 - ✓ 费用：不同问题用不同模型
 - ✓ 费用：优先选择工具和脚本，尽量形成Skills

2. 养龙虾就是养自己：

- 培养的是肉身：自己的AI认知和协同能力
- 培养的是分身：更强大更有精力的自己
- 类似于打（策略）游戏：需要天赋、有乐趣、易成瘾
- 知行合一做中学，先养一个

三个要素

1. 饲料最重要：

- 好饲料节省的是试错时间
- 自用和产品用可以区别对待
- Claude第1；GPT/Gemini第2；国产第3

2. 时间和脑子：

- 清楚自己的目标和场景
- 清楚该怎么一步一步做事

3. 产出和沉淀：

- Skill、Memory、Tools、Data

养虾的自我修养：AI Agent 进化指南

以“养虾”为喻，深度剖析自主 Agent 时代的构建、心态与核心要素，在低门槛高期望环境中实现业务自动化。

现状与心态 (STATUS & MINDSET)

自主 Agent 行业现状

- 没专家
- 变化快
- 平民化
- 门槛低
- 期望高



养虾即“养孩子”

必须投入充足的时间、费用与心力进行培育。



养虾即“养自己”

提升 AI 认知与协同能力，坚持知行合一，在实践中学习。

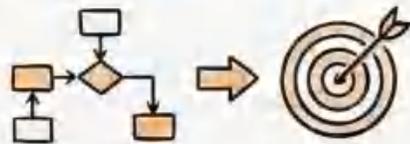
核心三要素 (CORE ELEMENTS)

要素一：饲料 (模型选择)

1. Claude (第一)
2. GPT 及 Gemini (第二)
3. 国产模型 (第三)



要素二：时间和脑子



必须明确目标与场景，脑子要清楚如何一步一步拆解任务。

要素三：产出和沉淀

SKILL (技能)

MEMORY (记忆)

TOOLS (工具)

DATA (数据)

核心资产包含 Skill、Memory、Tools 与 Data。

智能系统的 组合艺术

IT

- 命令、脚本/代码、软件工具、MCP/API
- 速度快、连续运行、确定性高、泛化能力弱
- 自动化白盒
- 代码逻辑为核心
- 边际效益高

AI

- 大模型、Agent
- 速度快、连续运行、确定性低、泛化能力强
- 智能化黑盒
- 模型计算为核心
- 有token成本，边际效益低

人

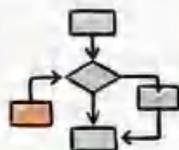
- 速度慢、间断运行、确定性低、泛化能力强
- 无边际效益，增加额外的沟通管理成本

智能系统的组合艺术：IT、AI 与人的深度协作

智能系统不仅是单一的技术，而是 IT、AI 与人三者结合的艺术。本图表对比了这三类要素在运行速度、确定性、泛化能力及核心逻辑上的本质区别，勾勒出完整系统的运行蓝图。

机器执行层：IT 与 AI 的双重特性

IT：高确定性的自动化白盒



CODE LOGIC
FAST & CONTINUOUS



HIGH CERTAINTY

以代码逻辑为核心，速度快且连续运行，虽泛化能力弱但边际效益极高。

AI：强泛化能力的智能化黑盒



HIGH GENERALIZATION

以模型计算为核心，兼具高速与强泛化性，但确定性较低且存在 Token 成本。



MODEL CALCULATION

TOKEN COST

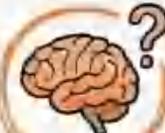


SPEED

GENERALIZATION

核心驱动层：人的角色与系统本质

人：系统中的高阶泛化节点



HIGH FLEXIBILITY

尽管速度较慢且间断运行，但人拥有极高的泛化能力和处理不确定性的潜力。

SLOW & INTERMITTENT

智能系统的本质：组合的艺术

IT (确定性)

AI (智能化)

人 (灵活性)



ART OF COMBINATION

智能系统是 IT (确定性)、AI (智能化) 与人 (灵活性) 的有机组合。

关键运行参数对比

| 维度 | IT | AI | 人 |
|----------|--------|--------|--------|
| 运行速度/持续性 | 快 / 连续 | 快 / 连续 | 慢 / 间断 |
| 确定性 | 高 | 低 | 低 |
| 泛化能力 | 弱 | 强 | 强 |

谢谢

肖睿

20260322@北大光华管理学院
