

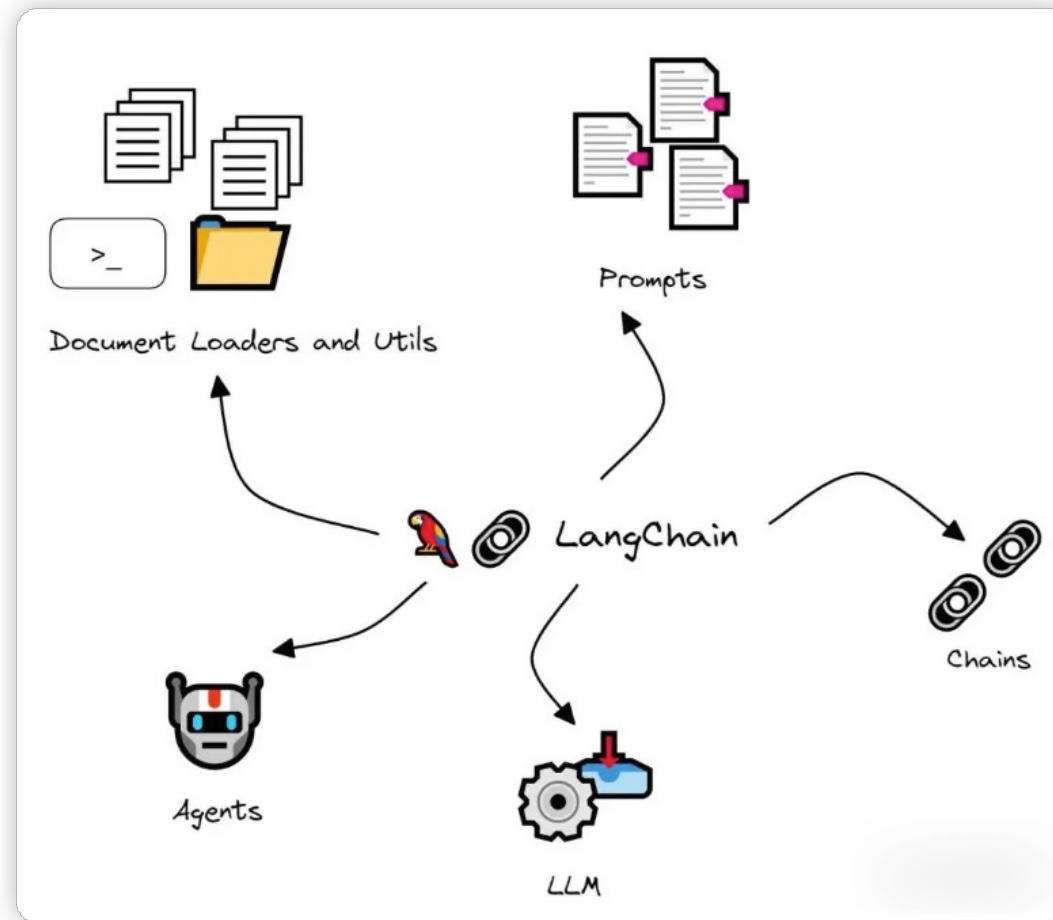
2-1 LangChain 概述

- 什么是LangChain
- LangChain的使用与安装
- LangChain 与OpenAI



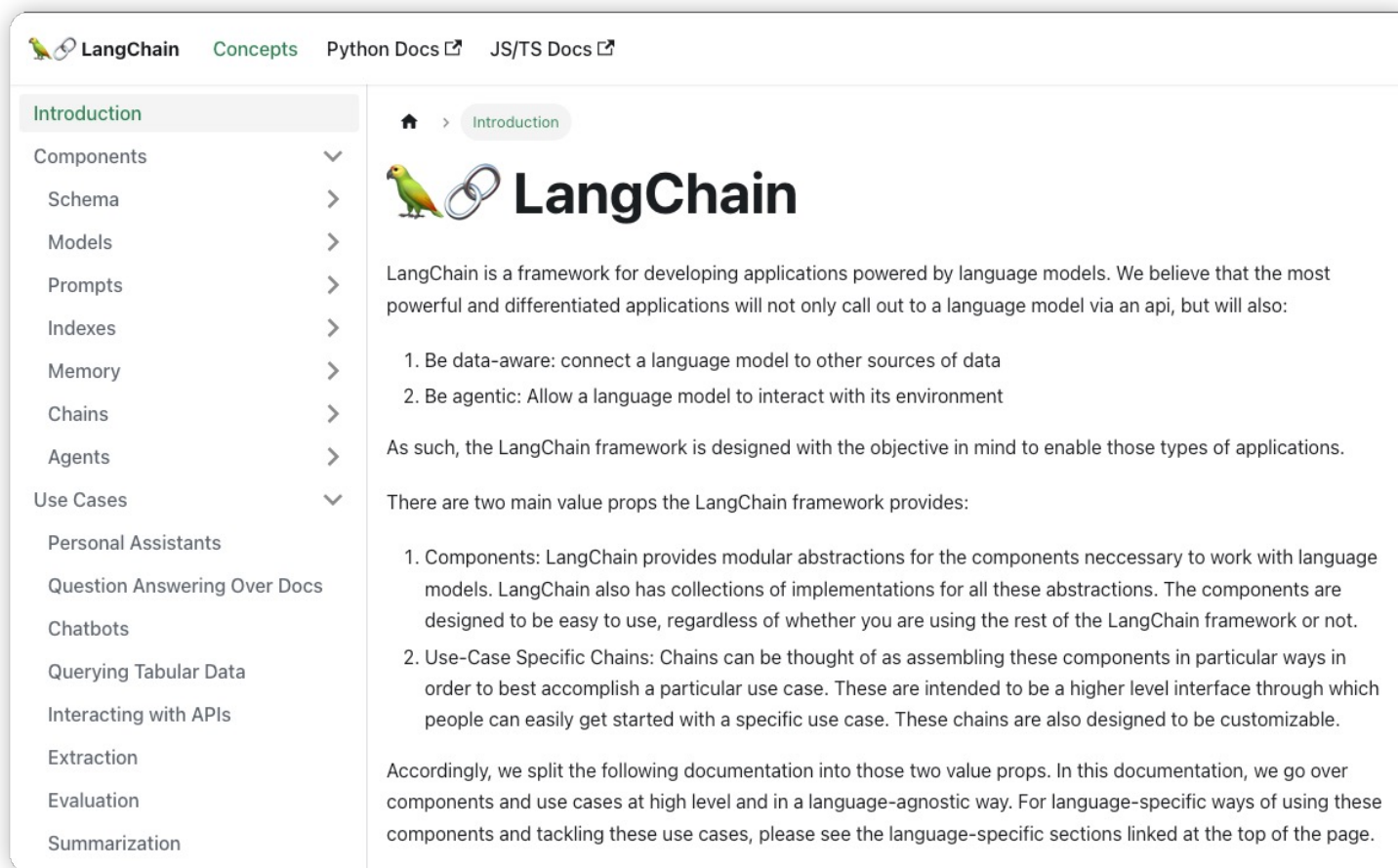
什么是LangChain

- 2022 年10 月 Harrison Chase 创造LangChain
- 集成：大语言模型、其他数据源（文档，数据库，应用程序、API）。
- 交互：让不同来源的数据进行交互
- 链：支持一连串的动作
- **围绕大模型的能力整活**



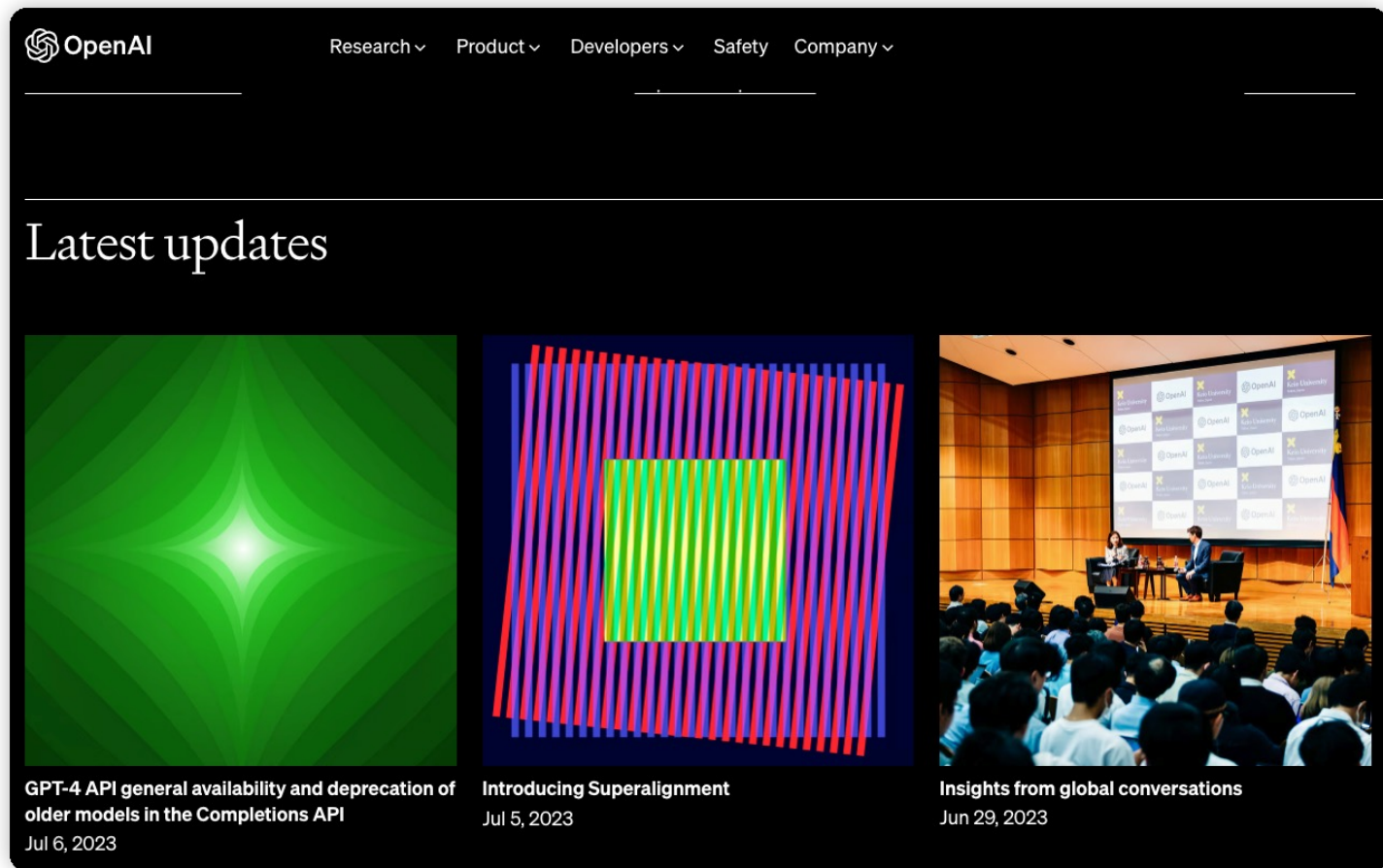
LangChain的安装与使用

- 官方网站 <https://langchain.com/>
- Python语言
- 安装: `pip install langchain`
- IDE:
 - CoLab (网络、Google、运算资源、集成)
 - anaconda (本地、网络依赖小)



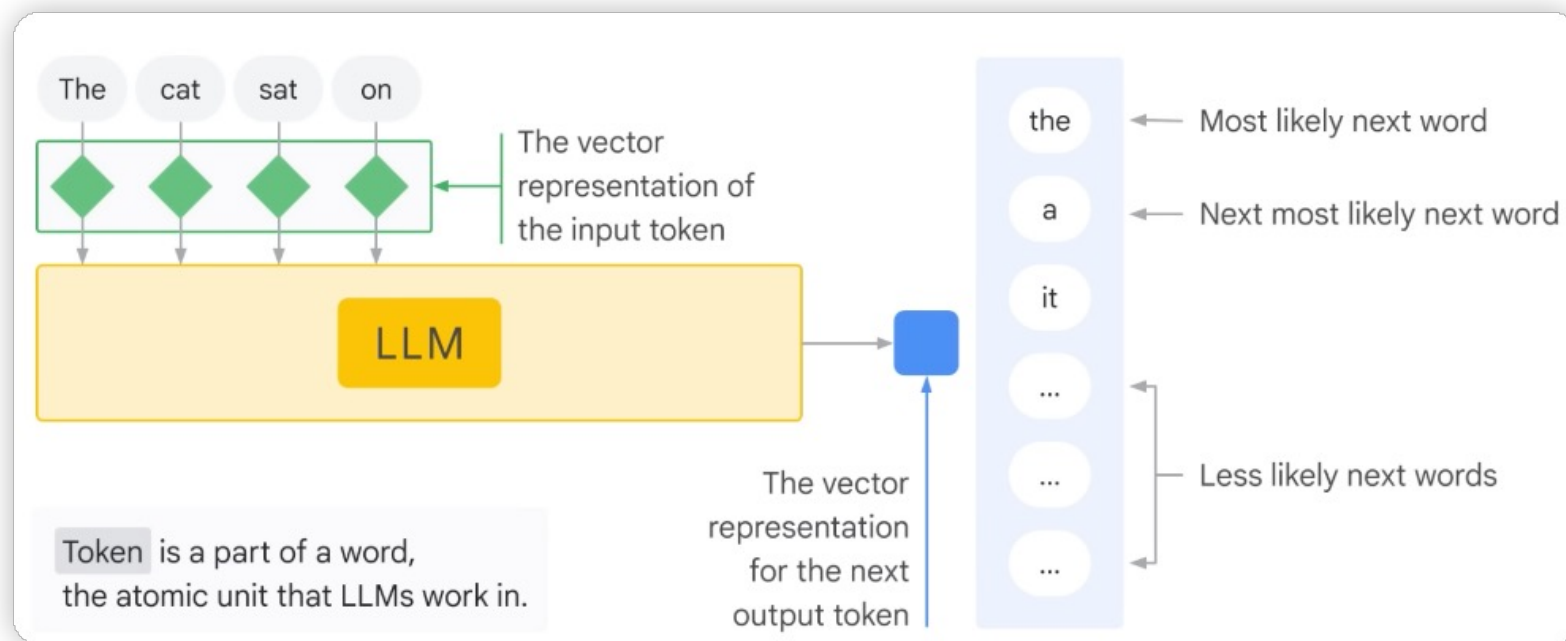
LangChain与OpenAI

- 使用OpenAI
 - 技术领先、易于使用
 - LLM场景广泛（语言、情感、翻译、问答）
 - 与LangChain集成良好
 - 社区良好，持续更新
- 申请OpenAI
 - OpenAI 账号
 - OpenAI plus（GPT 3.5->GPT 4）
 - OpenAI API 充值



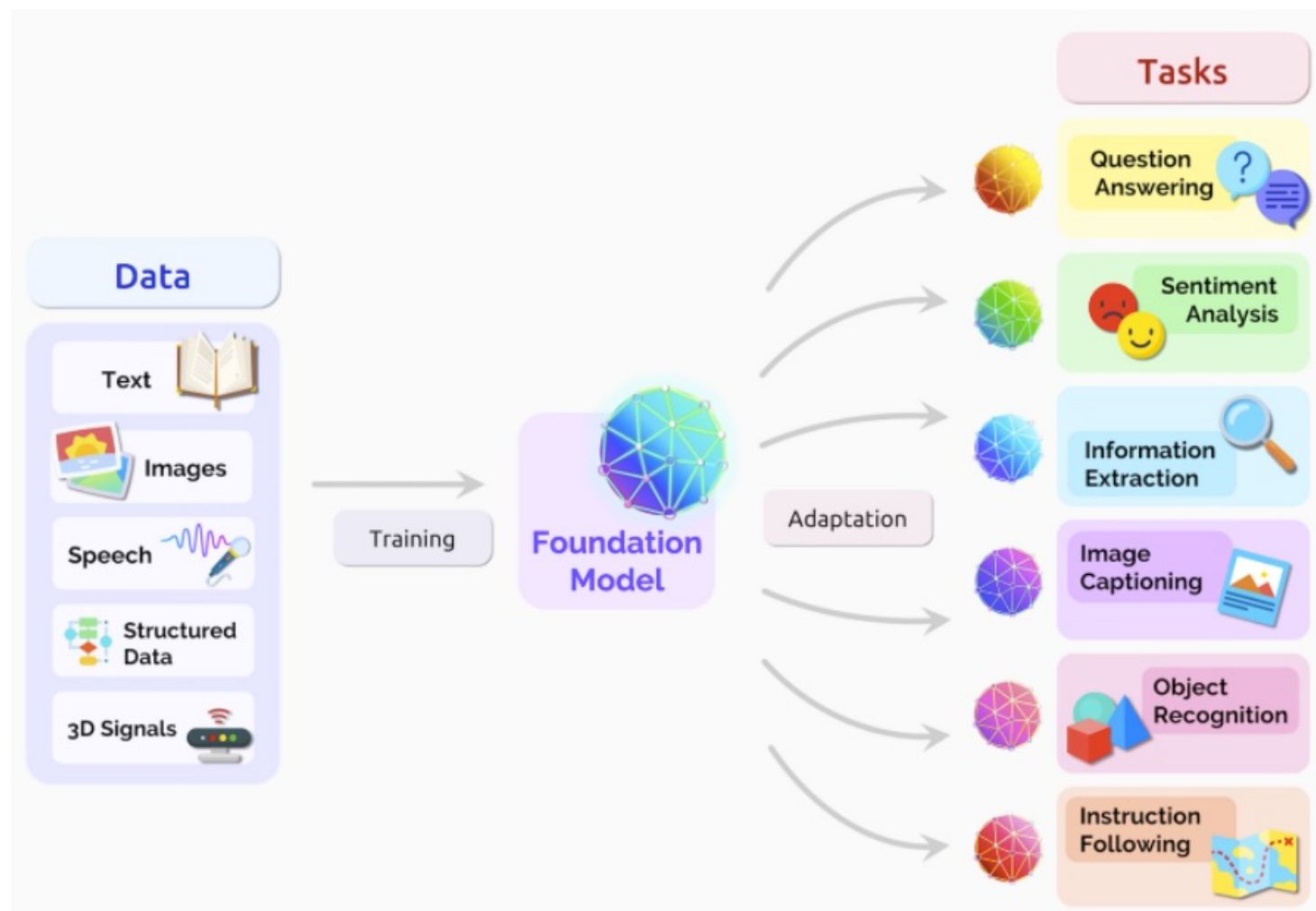
什么是大语言模型

- 理解、生成人类语言
- 总结，翻译，情感分析
- 简而言之就是**预测文本**



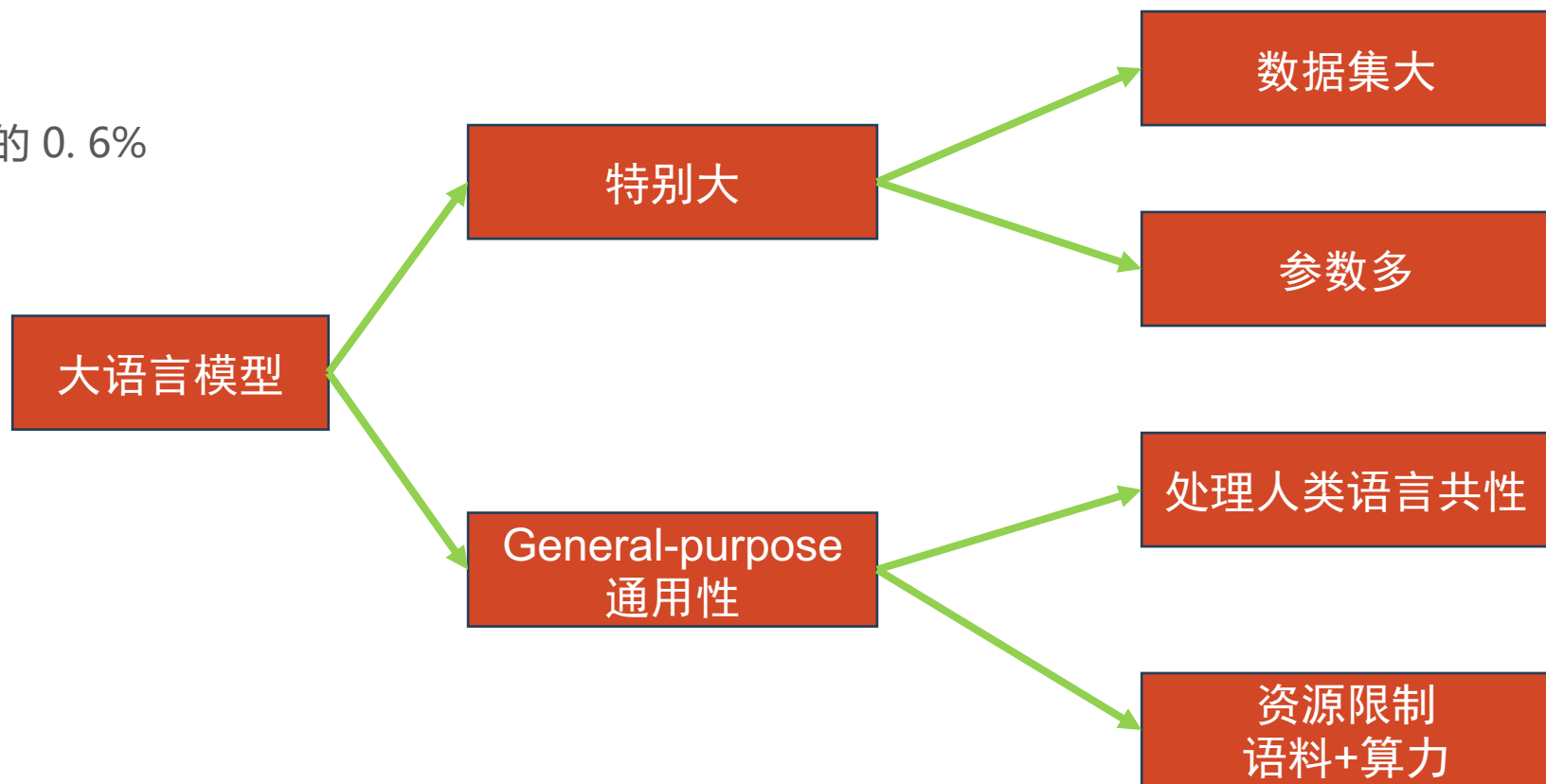
什么是大语言模型

- LLM 是Large Language Model的缩写。
- 人工智能模型
- 理解和生成人类语言
- 通过大量的文本的训练完成多种任务
- 文本总结、翻译、情感分析

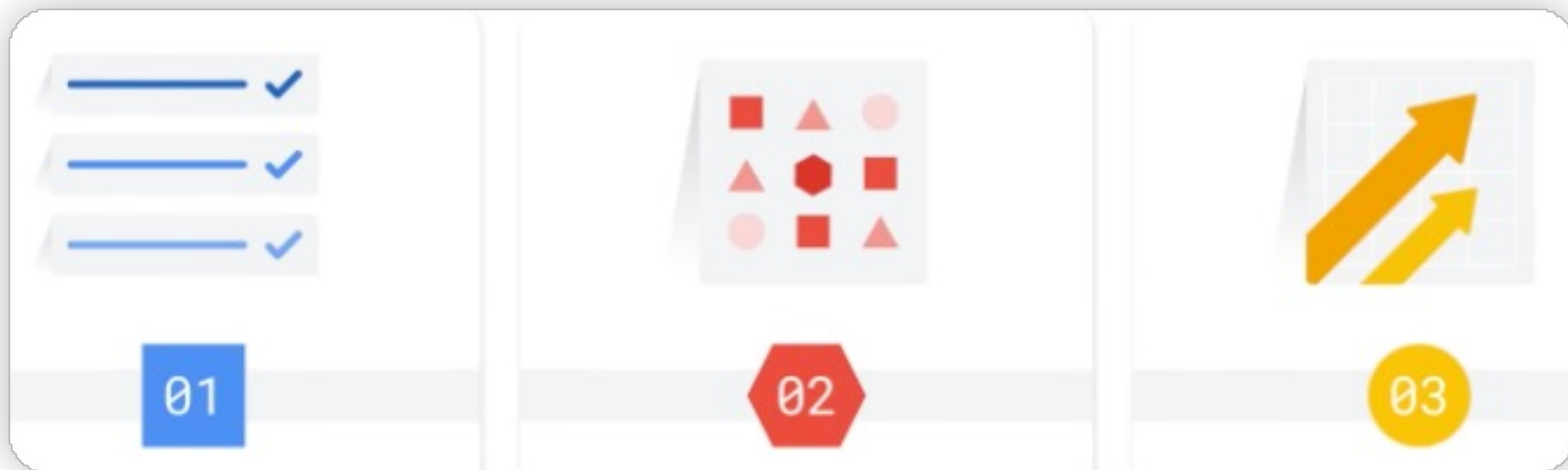


大语言模型的特点

- GPT 3 训练数据=45 TB
- 维基百科的数据=GPT3训练数据的 0.6%
- GPT3.5 参数=1750 亿
- 权重+偏执



大语言模型的优势



- 1 单一模型可用于不同任务
 - 通用能力
 - 通才+专才
- 2 少量数据进行微调
 - 相对预训练数据
- 3 数据集、参数越多，性能持续提升
 - 数据集提升
 - 参数提升
 - 性能提升

大语言模型的训练过程

