DeepSeek

每个人都可以读懂的大模型科普文章

大模型概念、技术与应用实践

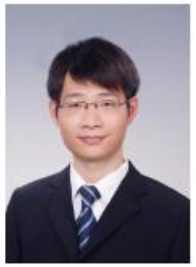
林子雨　副教授　　厦门大学

厦门大学大数据教学团队作品　2025年2月9日

厦门大学大数据教学团队

林子雨

夏小云


赖永炫

陶继平

谢怡

郑海山

蔡炳跃

苏淑文

郑宇辉

张琦

国内高校大数据教学的重要贡献者

团队负责人：林子雨 副教授
年轻力量：核心成员全部46周岁以下
结构合理：教学型、科研型、实验工程师
专注专业：从2013年至今，11年专注于大数据教学
团队特点：眼光前瞻、紧跟技术、创新实干、执行力强
影响力高：多项指标在国内高校大数据教学领域领先

- 教材数量
- 教材占有率
- MOOC课程学习人数
- 师资培养
- 教学研讨会
- 教学网站访问量
- 在线讲座观看人数
- ……

团队负责人林子雨

厦门大学计算机科学与技术系副教授

以第一作者编著出版15本大数据系列教材被国内1000余所高校采用

荣获“2022年福建省高等教育教学成果奖特等奖（个人排名第一）”

入选“2021年高校计算机专业优秀教师奖励计划”

2018年国家精品在线开放课程（独立主讲）

2020年国家级线上一流本科课程（独立主讲）

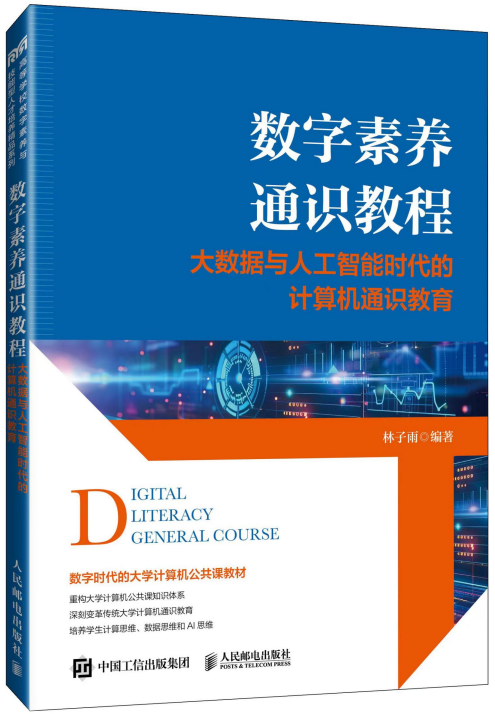
2021年国家级线上一流本科课程（独立主讲）

入选“2023年教育部国家智慧教育公共服务平台应用典型案例”

个人主页：https://dblab.xmu.edu.cn/post/linziyu/

本PPT节选自林子雨编著《数字素养通识教程》

林子雨 编著《数字素养通识教程——大数据与人工智能时代的计算机通识教育》
人民邮电出版社，2025年1月
ISBN:978-7-115-65946-0 定价：59.8元



面向大一新生的全校大学计算机公共课教材

教材官网：
https://dblab.xmu.edu.cn/post/digital-literacy/
教材官网提供讲义PPT、MOOC视频、案例视频、上机实验、教学大纲、课程思政案例、开学第一课讲座PPT等丰富的教学资源

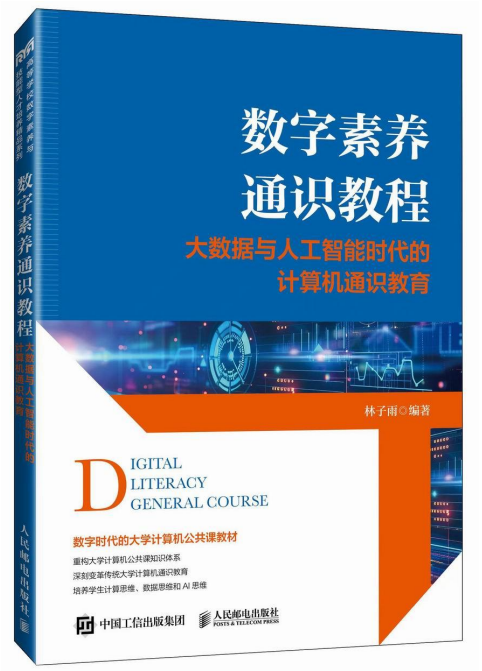
扫一扫访问教材官网

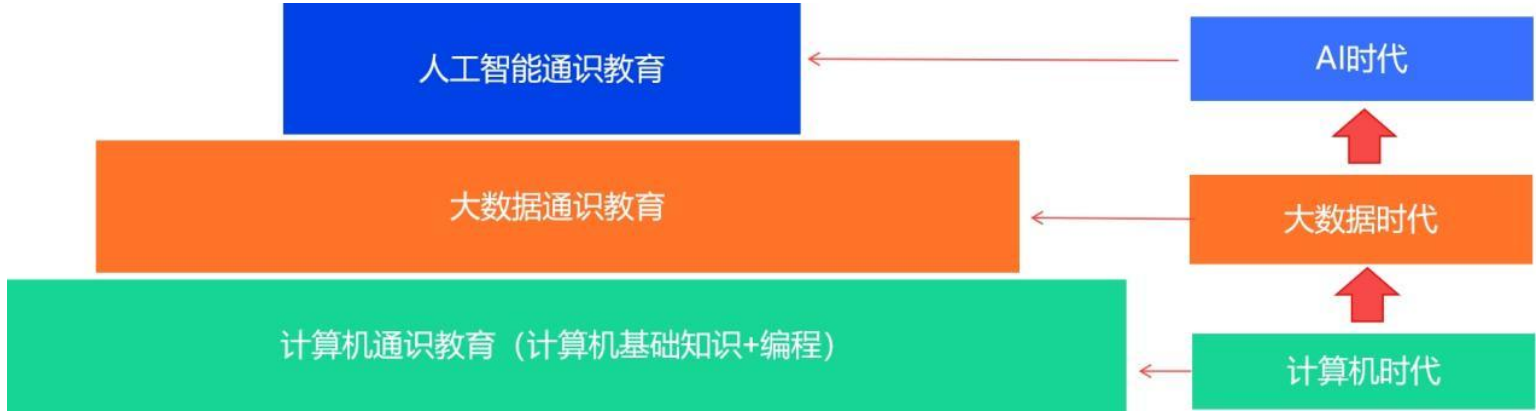
《数字素养通识教程》

林子雨编著《数字素养通识教程——大数据与人工智能时代的计算机通识教育》

人民邮电出版社 ISBN:978-7-115-65946-0 2025年1月第1版，定价：59.8元

教材官网：https://dblab.xmu.edu.cn/post/digital-literacy/





15年计算机教学生涯感悟升华，15本计算机畅销教材知识凝练
数字时代的大学计算机公共课教材，重构大学计算机公共课知识体系
深刻变革传统大学计算机通识教育，培养学生计算思维、数据思维和AI思维

《数字素养通识教程》

到B站观看林子雨主讲《数字素养通识教程》MOOC视频（1359分钟）
视频地址：https://www.bilibili.com/video/BV1XPf8YZE6M/

引言

在数字化浪潮汹涌澎湃的当下，大模型如同一颗璀璨新星，强势崛起并迅速成为科技领域的焦点。从最初的理论探索到如今在各个行业的广泛应用，大模型正以惊人的速度重塑着我们的生活与工作模式。它不仅是人工智能技术发展的重大突破，更是推动经济增长、提升社会治理效能、促进科技创新的关键力量。本报告《大模型概念、技术与应用实践》将深入剖析大模型的核心概念、原理特点以及丰富多元的应用实践案例，旨在让大家全面了解大模型这一前沿技术，明晰其在当下及未来发展中的重要地位与深远影响，共同探索如何借助大模型的力量推动社会各项事业迈向新的高度。

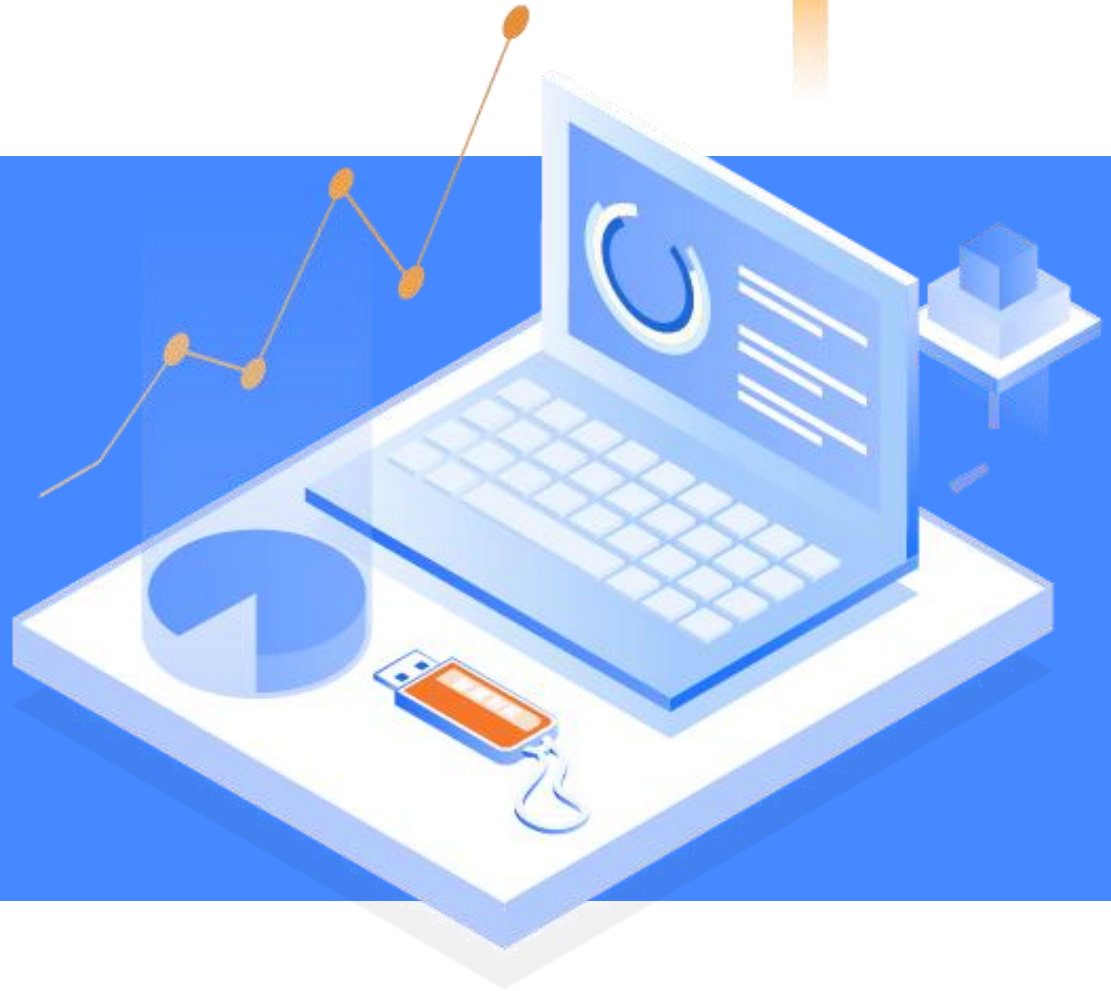
deepseek

目录

- 1. 人工智能发展简史
- 2. 人工智能思维
- 3. 大模型：人工智能的前沿
- 4. AIGC应用与实践

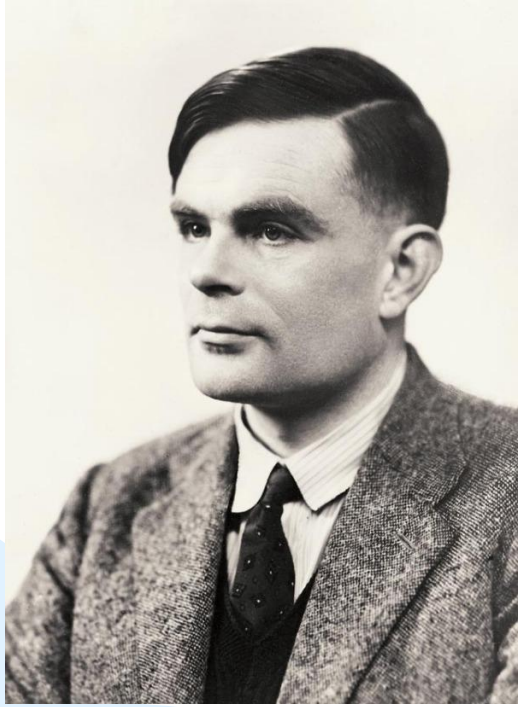
1. 人工智能发展简史

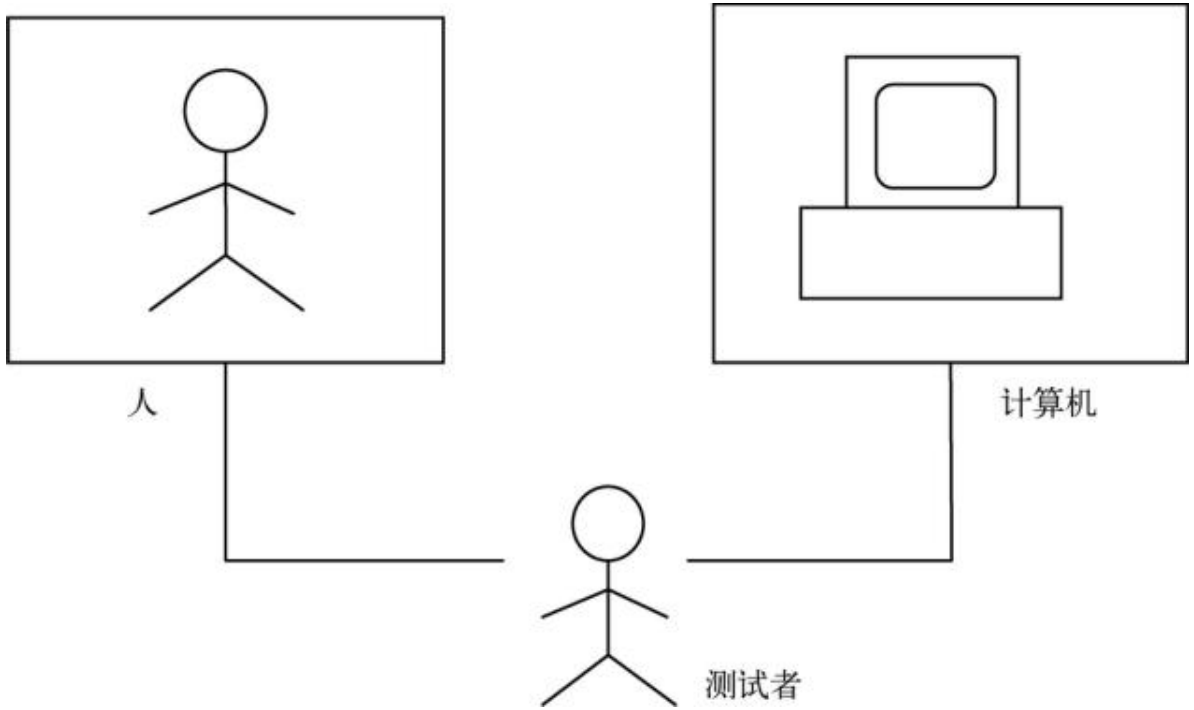
1.1 图灵测试
1.2 人工智能的诞生
1.3 人工智能的发展阶段
1.4 未来人工智能发展的五个阶段



1.1 图灵测试

1950年，“计算机之父”和“人工智能之父”艾伦·图灵（Alan M. Turing）发表了论文《计算机器与智能》，这篇论文被誉为人工智能科学的开山之作。在论文的开篇，图灵提出了一个引人深思的问题：“机器能思考吗？”。这个问题激发了人们无尽的想象，同时也奠定了人工智能的基本概念和雏形





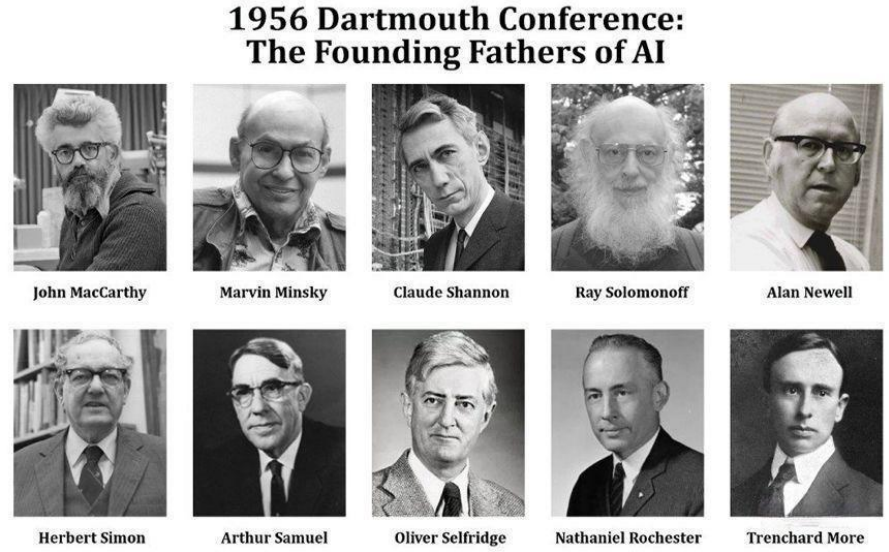
在这篇论文中，图灵提出了鉴别机器是否具有智能的方法，这就是人工智能领域著名的“图灵测试”。如图所示，其基本思想是测试者在与被测试者（一个人和一台机器）隔离的情况下，通过一些装置（如键盘）向被测试者随意提问。进行多次测试后，如果被测试者机器让平均每个测试者做出超过30%的误判，那么这台机器就通过了测试，并被认为具有人类智能

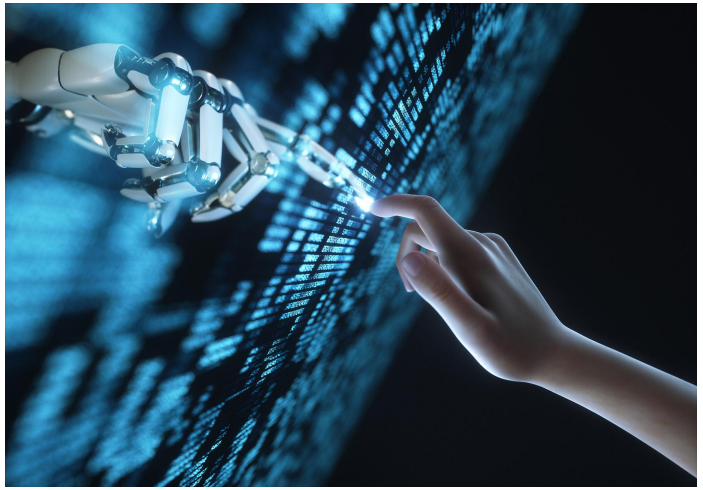
1.2 人工智能的诞生

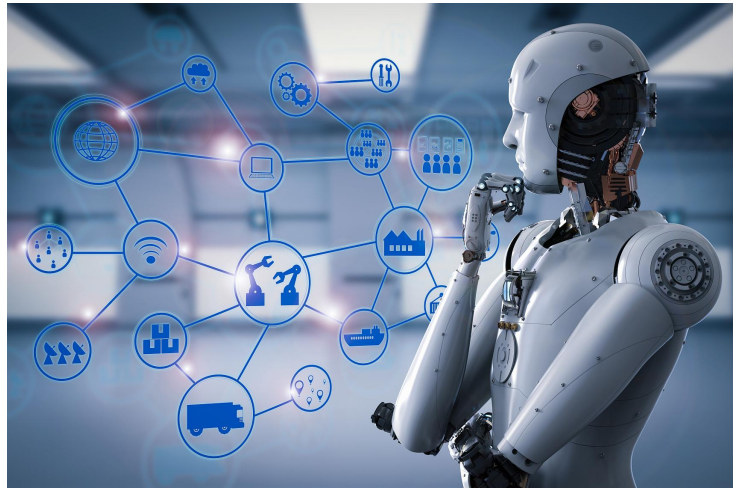
人工智能的诞生可以追溯到20世纪50年代。当时，计算机科学刚刚起步，人们开始尝试通过计算机程序来模拟人类的思维和行为。在这个背景下，一些杰出的科学家和工程师们开始研究如何使计算机具备更高级的功能

1956年8月，在美国达特茅斯学院举办的人工智能夏季研讨会，是人工智能领域具有里程碑意义的一次重要会议。这次会议汇聚了众多杰出的科学家和工程师，他们共同探讨和研究人工智能的发展和应用前景

这次会议的主题围绕着**人工智能的定义、研究方法和应用场景**展开。与会者们深入探讨了人工智能的基本概念、算法和技术，以及其在各个领域的应用潜力。他们共同认识到，人工智能的研究和发展将为人类带来巨大的变革和进步





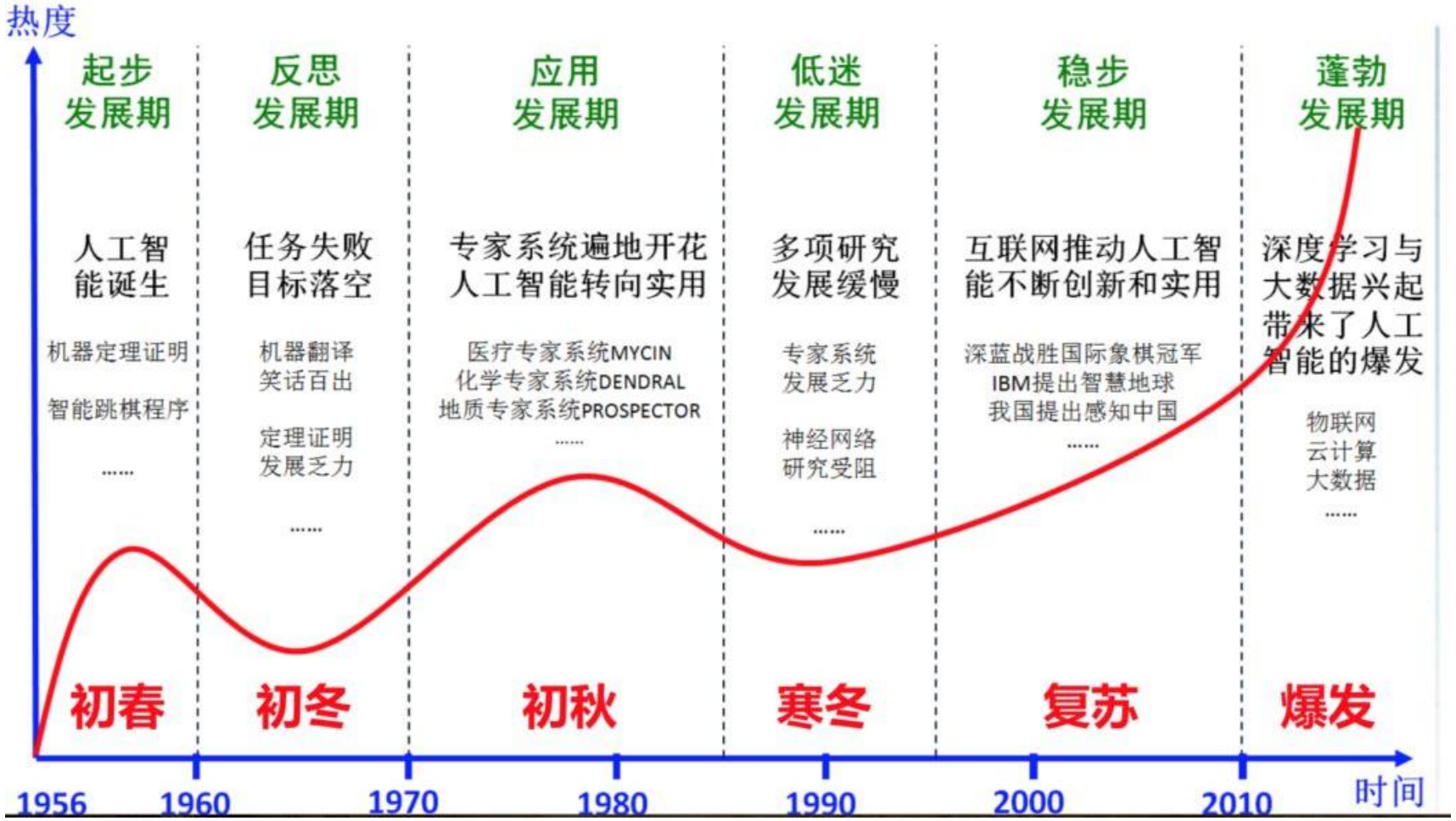


1.2 人工智能的诞生

在这次会议上，“人工智能”这个词汇被约翰.麦卡锡（John McCarthy）首次提出。与会者们不仅对人工智能的研究和应用前景进行了深入探讨，还提出了许多重要的观点和思路，为人工智能的发展奠定了基础。**这次会议的召开标志着人工智能作为一个独立学科的正式诞生，因此，达特茅斯会议被称为“人工智能的开端”，1956年也被称为“人工智能元年”**。这次会议不仅为人工智能的研究和发展奠定了基础，还为人类带来了巨大的变革和进步

1.3 人工智能的发展阶段

从1956年人工智能元年至今，人工智能的发展历程经历了漫长的岁月，大致可以划分为以下6个阶段



1.4 未来人工智能发展5个阶段

OpenAI 的 5 级 AGI 量表

1-5 级说明

L5：可以完成整个组织工作的人工智能。

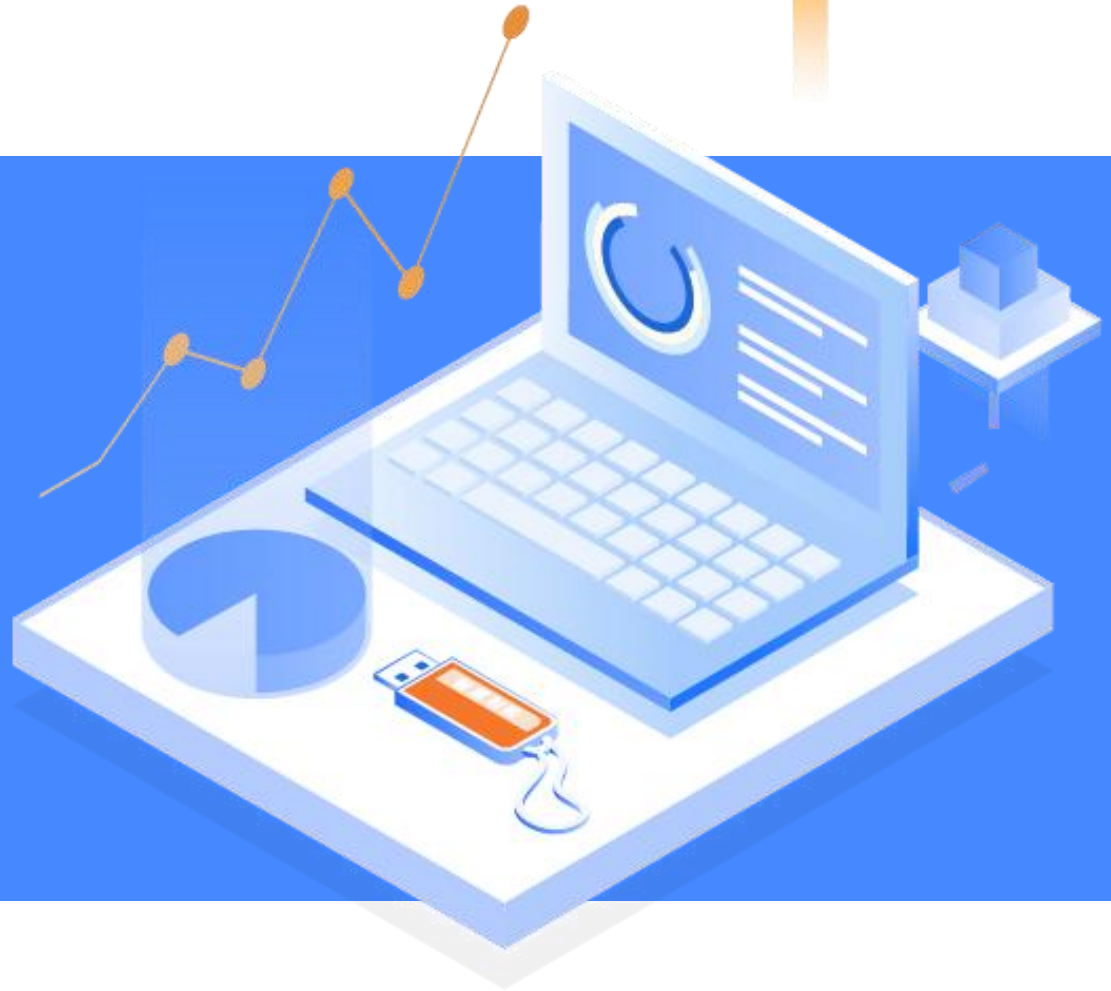
L4：可以创造新事物的人工智能。

L3：可以代表用户采取行动的人工智能代理。

L2：像博士一样解决问题，无需工具。

L1：当前AI，类似ChatGPT，可以与人类对话的人工智能。

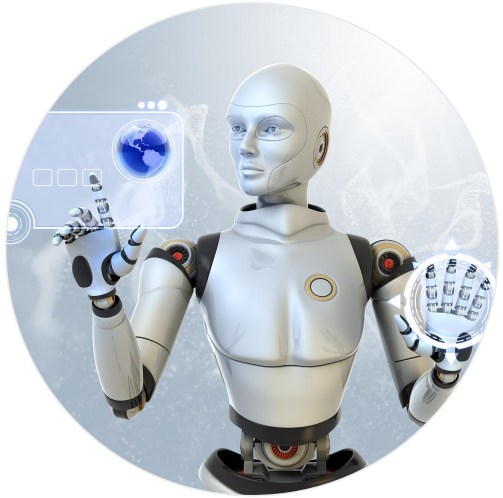
2. 人工智能思维



2. 人工智能思维

了解

每个人都应了解人工智能的基础运行模式



区分

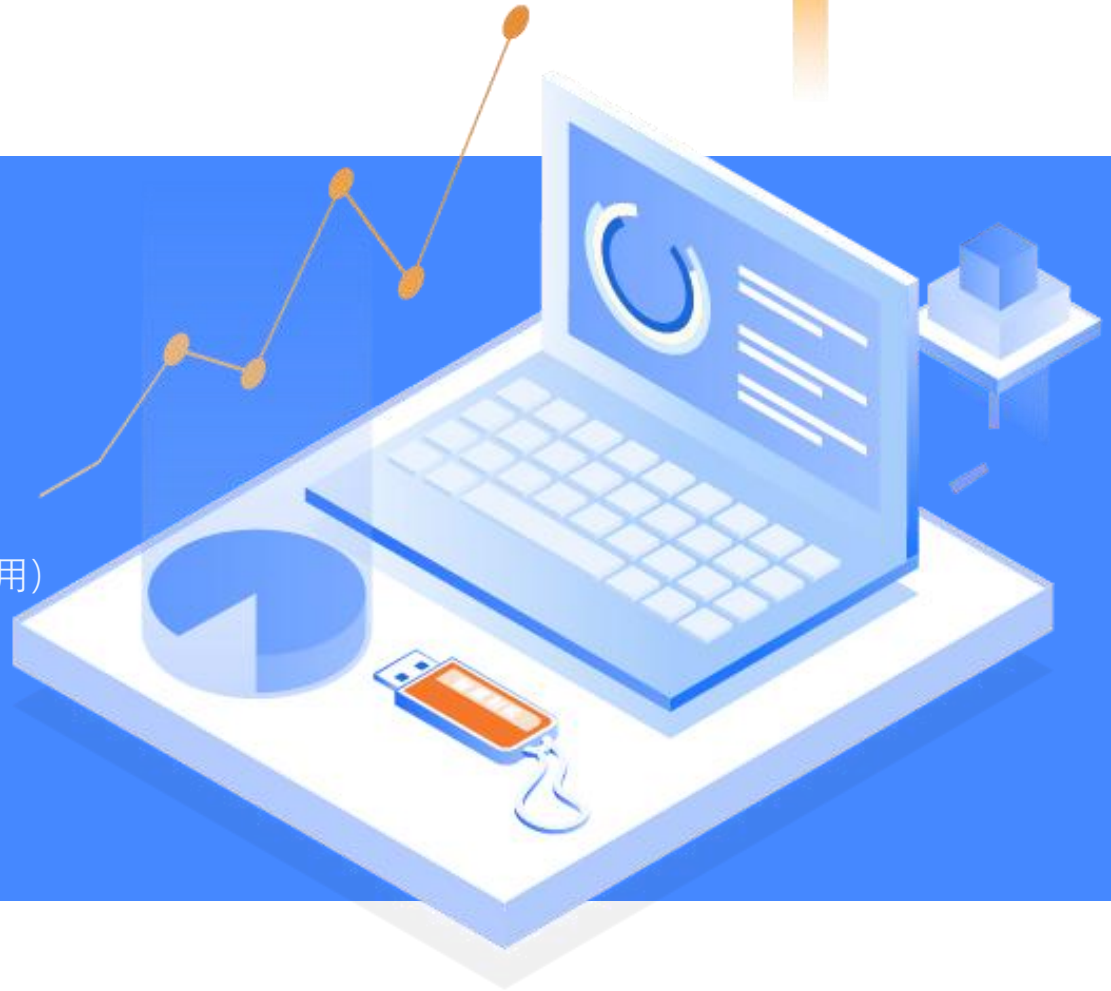
具备区分人的能力和机器的能力

协作

拥有和人工智能协作的能力，懂得如何运用人工智能

3.大模型：人工智能的前沿

3.1 大模型的概念
3.2 大模型的发展历程
3.3 人工智能与大模型的关系
3.4 大模型产品
3.5 大模型原理
3.6 大模型特点
3.7 大模型分类
3.8 大模型应用领域（在各个行业的应用）
3.9 大模型对工作和生活的影响
3.10 本地部署大模型
3.11 基于大模型的智能体



3.1 大模型的概念

大模型通常指的是大规模的人工智能模型，是一种基于深度学习技术，具有海量参数、强大的学习能力和泛化能力，能够处理和生成多种类型数据的人工智能模型

通常说的大模型的“大”的特点体现在：**参数数量庞大、训练数据量大、计算资源需求高**

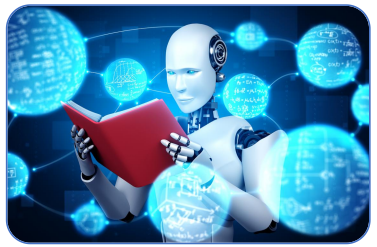
2020年，OpenAI公司推出了GPT-3，模型参数规模达到了1750亿，2023年3月发布的GPT-4的参数规模是GPT-3的10倍以上，达到1.8万亿，2021年11月阿里推出的M6 模型的参数量达10万亿

3.1 大模型的概念

大模型的设计和训练旨在提供更强大、更准确的模型性能，以应对更复杂、更庞大的数据集或任务。大模型通常能够学习到更细微的模式和规律，具有更强的泛化能力和表达能力

上下文理解能力

大模型具有更强的上下文理解能力，能够理解更复杂的语意和语境。这使得它们能够产生更准确、更连贯的回答

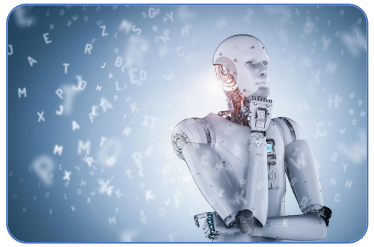


语言生成能力

大模型可以生成更自然、更流利的语言，减少了生成输出时呈现的错误或令人困惑的问题

学习能力强

大模型可以从大量的数据中学习，并利用学到的知识和模式来提供更精准的答案和预测。这使得它们在解决复杂问题和应对新的场景时表现更加出色

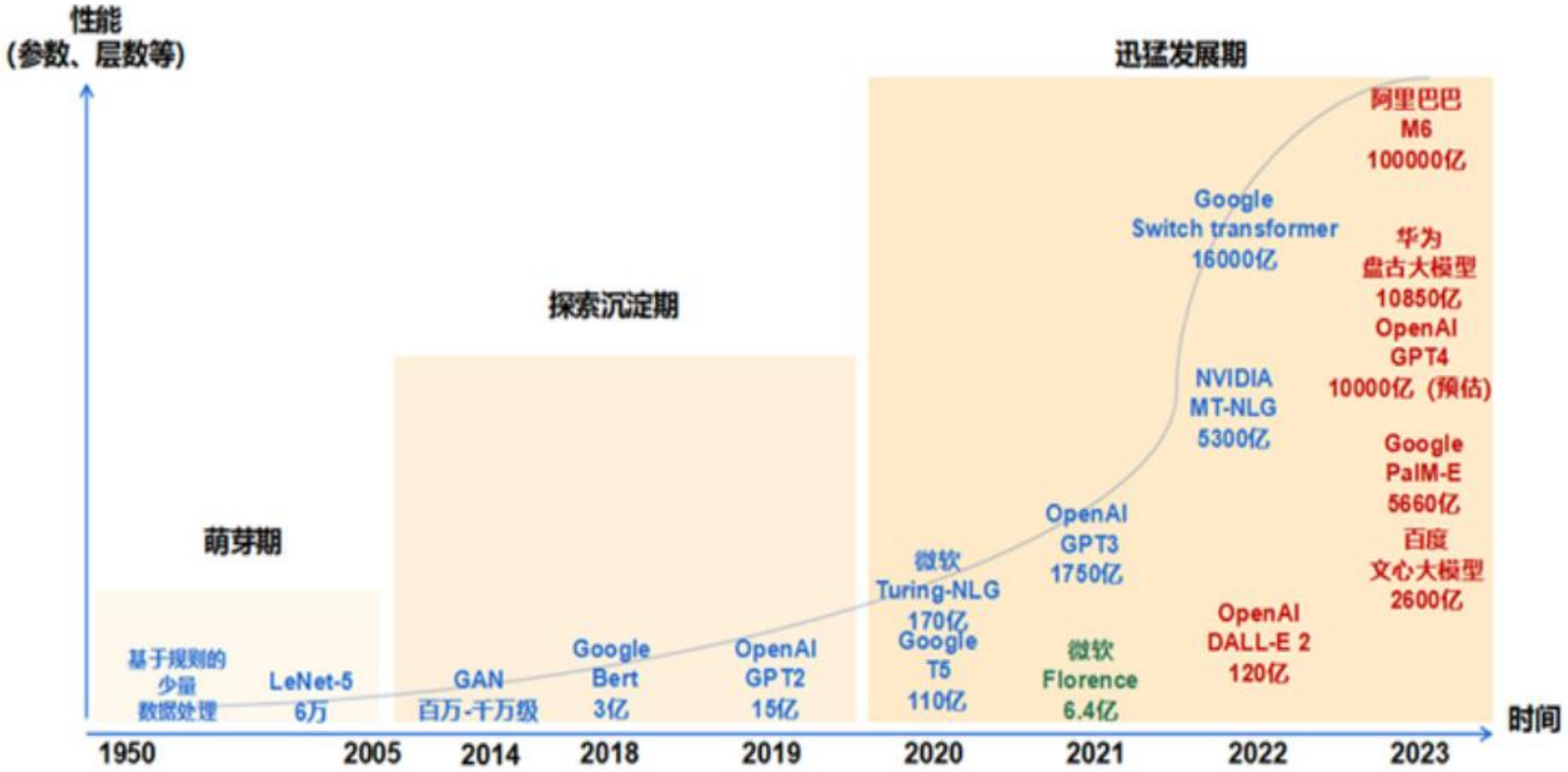


可迁移性高

学习到的知识和能力可以在不同的任务和领域中迁移和应用。这意味着一次训练就可以将模型应用于多种任务，无需重新训练

3.2 大模型的发展历程

大模型发展历经三个阶段，分别是萌芽期、沉淀期和爆发期

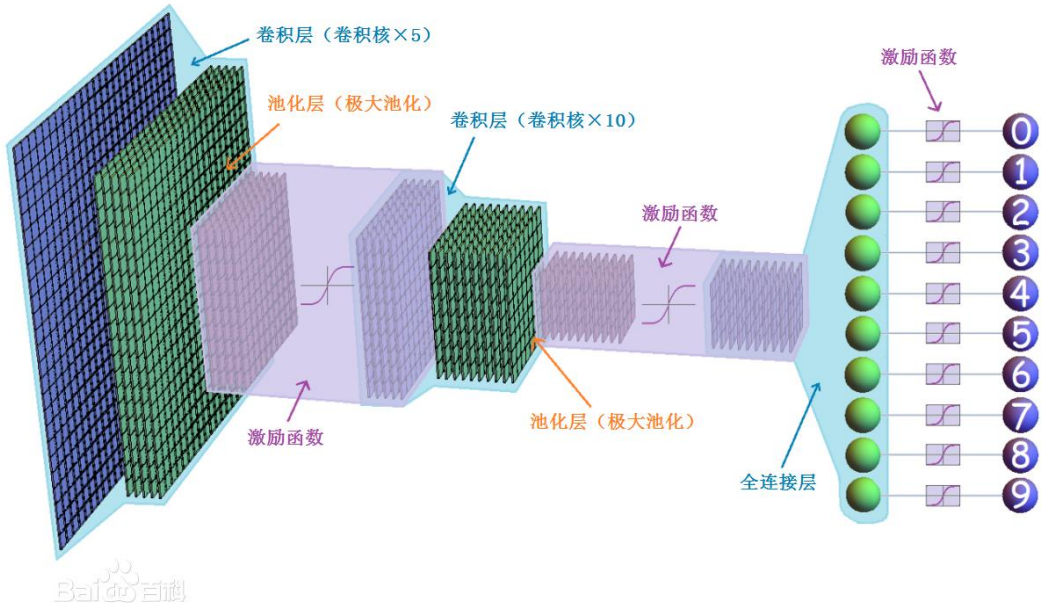


3.2 大模型的发展历程

■ 萌芽期（1950-2005）

这是一个**以CNN（Convolutional Neural Networks，卷积神经网络）为代表的传统神经网络模型阶段**

- 1956年，从计算机专家约翰·麦卡锡提出“人工智能”概念开始，AI发展由最开始基于小规模专家知识逐步发展为基于机器学习
- 1980年，卷积神经网络的雏形CNN诞生
- 1998年，现代卷积神经网络的基本结构LeNet-5诞生，机器学习方法由早期基于浅层机器学习的模型，变为了基于深度学习的模型，为自然语言生成、计算机视觉等领域的深入研究奠定了基础，对后续深度学习框架的迭代及大模型发展具有开创性的意义



3.2 大模型的发展历程

■ 沉淀期（2006-2019）

这是一个**以Transformer为代表的全新神经网络模型阶段**

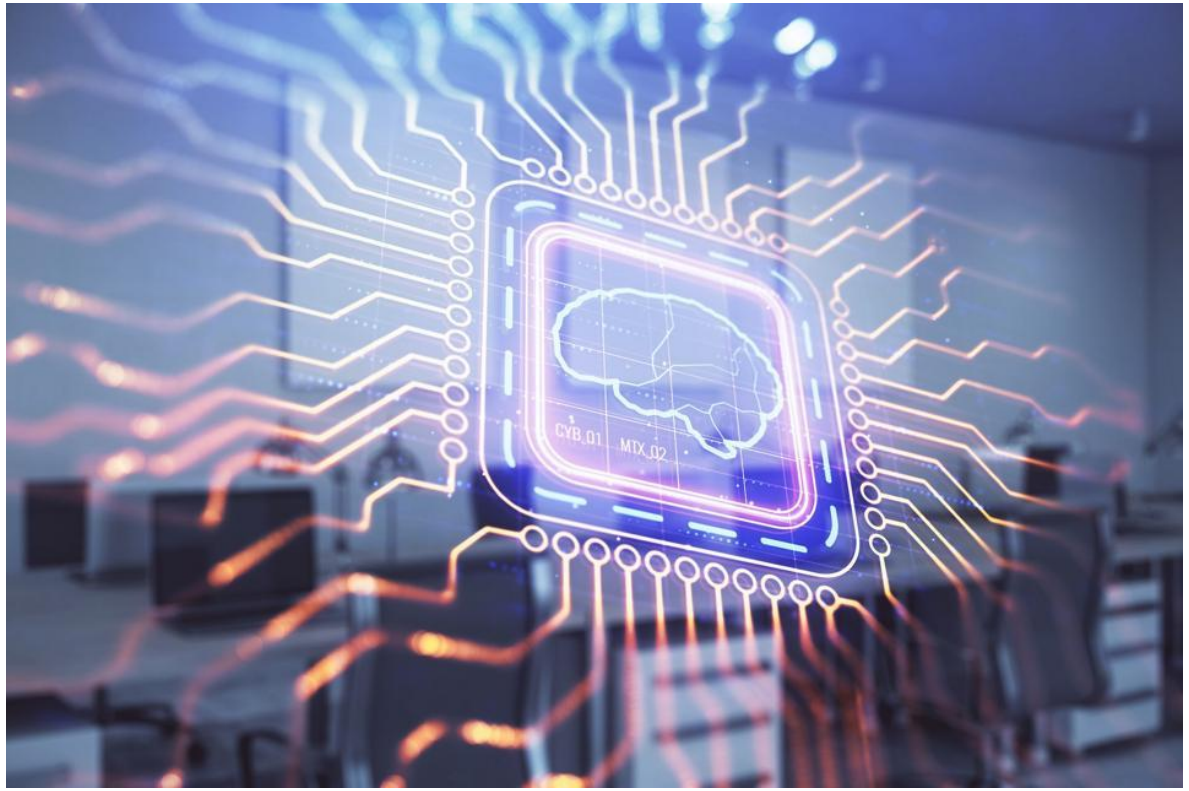
2013年，自然语言处理模型 Word2Vec诞生，首次提出将单词转换为向量的“词向量模型”，以便计算机更好地理解和处理文本数据。

2014年，被誉为21世纪最强大算法模型之一的GAN（Generative Adversarial Networks，对抗式生成网络）诞生，标志着深度学习进入了生成模型研究的新阶段

2017年，Google颠覆性地提出了基于自注意力机制的神经网络结构——Transformer架构，奠定了大模型预训练算法架构的基础

2018年，OpenAI基于Transformer架构发布了GPT-1大模型，意味着预训练大模型成为自然语言处理领域的主流，其中，GPT的英文全称是Generative Pre-Trained Transformer，是一种基于互联网的、可用数据来训练的、文本生成的深度学习模型

2019年，OpenAI发布了GPT-2



3.2 大模型的发展历程

■ 爆发期（2020-至今）

这是一个**以GPT为代表的预训练大模型阶段**

2020年6月，OpenAI公司推出了GPT-3，模型参数规模达到了1750亿，成为当时最大的语言模型，并且在零样本学习任务上实现了巨大性能提升。随后，更多策略如基于人类反馈的强化学习（RLHF，Reinforcement Learning from Human Feedback）、代码预训练、指令微调等开始出现，被用于进一步提高推理能力和任务泛化

2022年11月，搭载了GPT3.5的ChatGPT（Chat Generative Pre-trained Transformer）横空出世，凭借逼真的自然语言交互与多场景内容生成能力，迅速引爆互联网，在全球范围内引起轰动，使得大模型的概念迅速进入普通大众的视野。ChatGPT是人工智能技术驱动的自然语言处理工具，它能够通过理解和学习人类的语言来进行对话，还能根据聊天的上下文进行互动，真正像人类一样来聊天交流，甚至能完成撰写邮件、视频脚本、文案、翻译、代码，写论文等任务

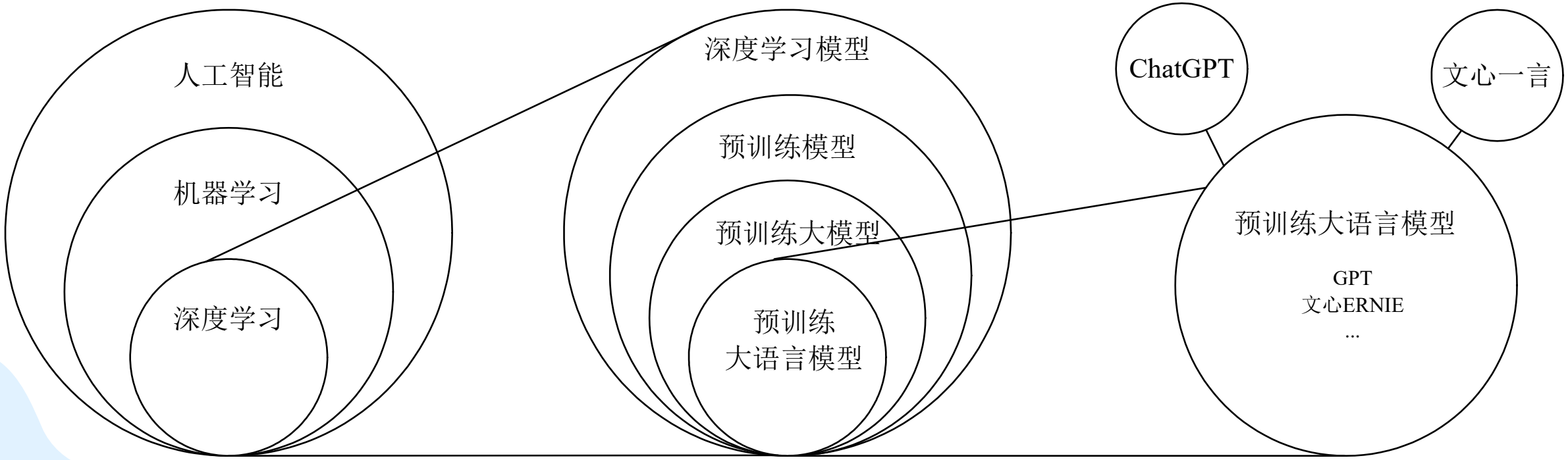
3.2 大模型的发展历程

■ 爆发期（2020-至今）

OpenAI在2023年3月发布了GPT-4，它是一个多模态大模型（接受图像和文本输入，生成文本）。相比上一代的GPT-3，GPT-4可以更准确地解决难题，具有更广泛的常识和解决问题的能力。2023年12月，谷歌发布大模型Gemini，它可以同时识别文本、图像、音频、视频和代码五种类型信息，还可以理解并生成主流编程语言（如Python、Java、C++）的高质量代码，并拥有全面的安全性评估。2024年12月，DeepSeek迅速崛起，震撼全球，使得人工智能进入“普惠”时代

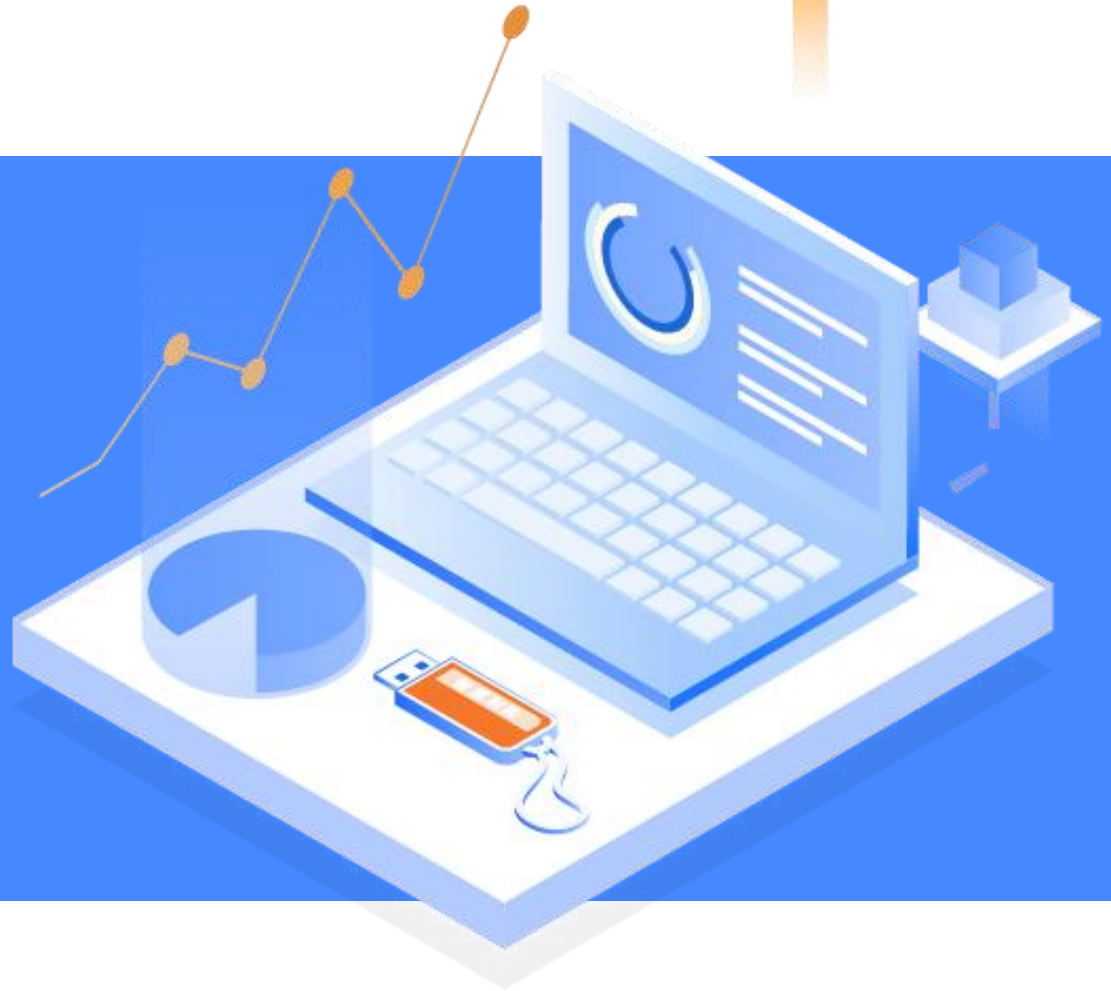
3.3 人工智能与大模型的关系

人工智能包含了机器学习，机器学习包含了深度学习，深度学习可以采用不同的模型，其中一种模型是预训练模型，**预训练模型包含了预训练大模型（可以简称为“大模型”），预训练大模型包含了预训练大语言模型（可以简称为“大语言模型”）**，预训练大语言模型的典型代表包括OpenAI的GPT和百度的文心ERNIE，ChatGPT是基于GPT开发的大模型产品，文心一言是基于文心ERNIE开发的大模型产品



3.4 大模型产品

3.4.1 国外的大模型产品
3.4.2 国内的大模型产品

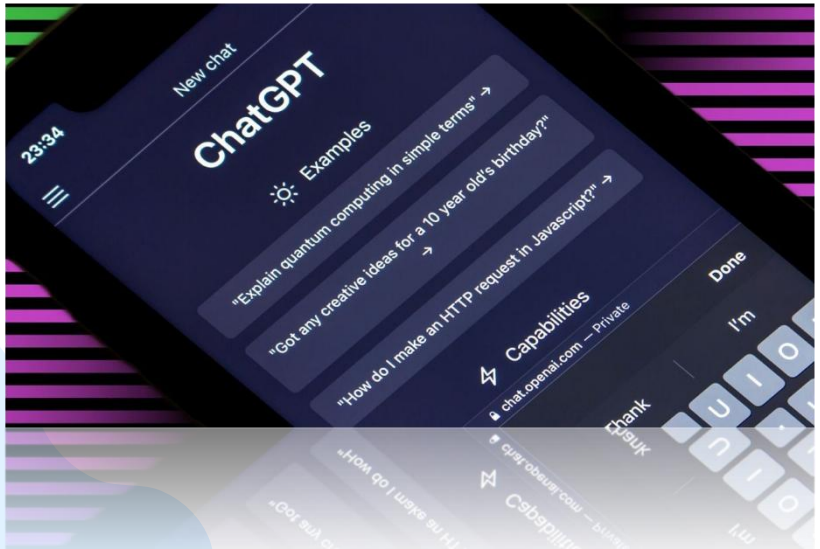


3.4.1 国外的大模型产品

■ ChatGPT

ChatGPT是一种**由OpenAI训练的大语言模型**。它是基于Transformer架构，经过大量文本数据训练而成，能够生成自然、流畅的语言，并具备回答问题、生成文本、语言翻译等多种功能

ChatGPT的应用范围广泛，**可以用于客服、问答系统、对话生成、文本生成等领域**。它能够理解人类语言，并能够回答各种问题，提供相关的知识和信息。与其他聊天机器人相比，ChatGPT具备更强的语言理解和生成能力，能够更自然地与人类交流，并且能够更好地适应不同的领域和场景。ChatGPT的训练数据来自互联网上的大量文本，因此，它能够涵盖多种语言风格和文化背景



3.4.1 国外的大模型产品

■ Gemini

Gemini是谷歌发布的大模型，它能够**同时处理多种类型的数据和任务，覆盖文本、图像、音频、视频等多个领域**。Gemini采用了全新的架构，将多模态编码器和多模态解码器两个主要组件结合在一起，以提供最佳结果

Gemini包括三种不同规模的模型：Gemini Ultra、Gemini Pro和Gemini Nano，适用于不同任务和设备。2023年12月6日，Gemini的初始版本已在Bard中提供，开发人员版本可通过Google Cloud的API获得。Gemini可以应用于Bard和Pixel 8 Pro智能手机。Gemini的应用范围广泛，包括**问题回答、摘要生成、翻译、字幕生成、情感分析**等任务。然而，**由于其复杂性和黑箱性质，Gemini的可解释性仍然是一个挑战**

3.4.1 国外的大模型产品

■ Sora

2024年2月16日，OpenAI再次震撼全球科技界，发布了名为Sora的文本生成视频大模型，只需输入文本就能自动生成视频。这一技术的诞生，**不仅标志着人工智能在视频生成领域的重大突破，更引发了关于人工智能发展对人类未来影响的深刻思考**。随着Sora的发布，人工智能似乎正式踏入了通用人工智能（AGI：Artificial General Intelligence）的时代。**AGI是指能够像人类一样进行各种智能活动的机器智能，包括理解语言、识别图像、进行复杂推理等**。Sora大模型能够直接输出长达60秒的视频，并且视频中包含了高度细致的背景、复杂的多角度镜头，以及富有情感的多个角色。这种能力已经超越了简单的图像或文本生成，开始触及到视频这一更加复杂和动态的媒介。这**意味着人工智能不仅在处理静态信息上越来越强大，而且在动态内容的创造上也展现出了惊人的潜力**

3.4.1 国外的大模型产品

■ Sora

右图是Sora根据文本自动生成的视频画面，一位戴着墨镜、穿着皮衣的时尚女子走在雨后夜晚的东京市区街道上，抹了鲜艳唇彩的唇角微微翘起，即便带着墨镜也能看到她的微笑，地面的积水映出了她的身影和灯红酒绿的霓虹灯，热闹非凡的唐人街正在进行舞龙表演，熙熙攘攘的人群目光都聚焦在跃动的彩龙身上，整个环境的喜庆氛围仿佛令人身临其境

提示词译文：

一位时尚女性走在布满温暖霓虹灯和动画城市标牌的东京街道上。 她穿着黑色皮夹克、红色长裙和黑色靴子，拎着黑色钱包。 她戴着太阳镜，涂着红色口红。 她走路自信又随意。 街道潮湿且反光，在彩色灯光的照射下形成镜面效果。 许多行人走来走去。

Sora生成的视频：

3.4.1 国外的大模型产品

■ OpenAI o3

2024年12月20日，OpenAI发布推理模型o3，无论在软件工程、编写代码，还是竞赛数学、掌握人类博士级别的自然科学知识能力方面，o3都达到了很高的水平

3.4.2 国内的大模型产品

2025年1月国内大模型排行榜

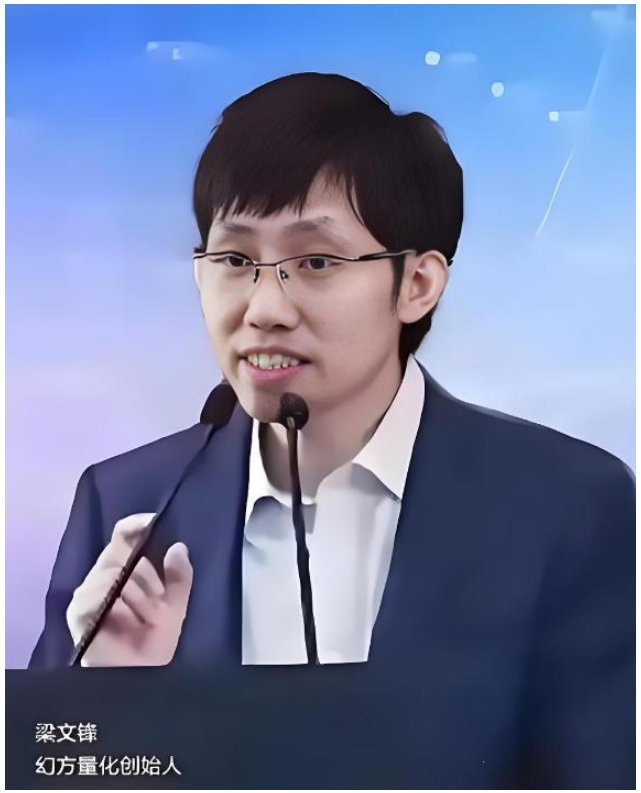
大模型	图标	指标排名
DeepSeek		能力测评第一
豆包		用户数量第一
Kimi		文本处理第一
即梦AI		作图能力第一
通义万相		视频生成第一
智谱清言		文档归纳第一

3.4.2 国内的大模型产品

■ DeepSeek（深度求索）

2024年12月26日，杭州一家名为“深度求索”（DeepSeek）的中国初创公司，发布了全新一代大模型DeepSeek-V3。在多个基准测试中，DeepSeek-V3的性能均超越了其他开源模型，甚至与顶尖的闭源大模型GPT-4o不相上下，尤其在数学推理上，DeepSeek-V3更是遥遥领先。DeepSeek-V3以多项开创性技术，大幅提升了模型的性能和训练效率。DeepSeek-V3在性能比肩GPT-4o的同时，研发却只花了558万美元，训练成本不到后者的二十分之一。因为表现太过优越，DeepSeek在硅谷被誉为“来自东方的神秘力量”。

2025年1月20日，DeepSeek-R1正式发布，拥有卓越的性能，在数学、代码和推理任务上可与OpenAI o1媲美。



DeepSeek创始人
梁文峰

3.4.2 国内的大模型产品

■ 通义千问

通义千问是阿里云推出的一个**超大规模的语言模型**，它具备**多轮对话、文案创作、逻辑推理、多模态理解、多语言支持**的能力。通义千问这个名字有“通义”和“千问”两层含义，“通义”表示这个模型能够理解各种语言的含义，“千问”则表示这个模型能够回答各种问题。通义千问基于深度学习技术，通过对大量文本数据进行训练，从而具备了强大的语言理解和生成能力。它能够理解自然语言，并能够生成自然语言文本

同时，通义千问还具备**多模态理解能力，能够处理图像、音频等多种类型的数据**。通义千问的应用范围非常广泛，可以**应用于智能客服、智能家居、移动应用等多个领域**。它可以与用户进行自然语言交互，帮助用户解决各种问题，提供相关的知识和信息。同时，通义千问还可以**与各种设备和应用进行集成**，为用户提供更加便捷的服务

3.4.2 国内的大模型产品

■ 字节跳动豆包

豆包是字节跳动基于云雀模型开发的 AI，能理解你的需求并生成高质量回应。它知识储备丰富，涵盖历史、科学、技术等众多领域，无论是日常问题咨询，还是深入学术探讨，都能提供准确全面的信息。同时，具备出色的文本创作能力，能撰写故事、诗歌、文案等各类体裁。并且擅长语言交互，交流自然流畅，就像身边的知心伙伴，耐心倾听并给予恰当反馈。

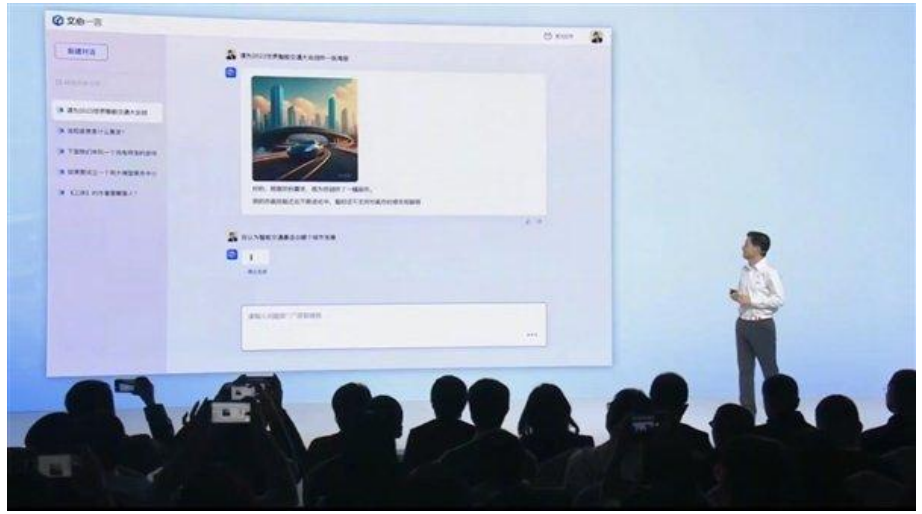
3.4.2 国内的大模型产品

■ 文心一言

文心一言是由百度研发的知识增强大模型，能够与人对话互动、回答问题、协助创作，高效便捷地帮助人们获取信息、知识和灵感

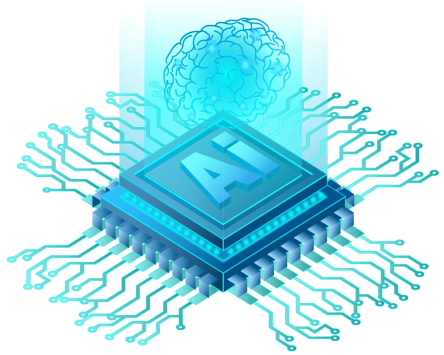
文心一言基于飞桨深度学习平台和文心知识增强大模型，持续从海量数据和大规模知识中融合学习，具备知识增强、检索增强和对话增强的技术特色。

文心一言**具有广泛的应用场景**，例如智能客服、智能家居、移动应用等领域。它可以与用户进行自然语言交互，帮助用户解决各种问题，提供相关的知识和信息



3.5 大模型的基本原理

大模型是基于Transformer架构的，这种架构是一种专门用于自然语言处理的“编码-解码器”架构。在训练过程中，大模型将输入的单词以向量的形式传递给神经网络，然后通过网络的编码解码以及自注意力机制，建立起每个单词之间联系的权重。大模型的核心能力在于将输入的每句话中的每个单词与已经编码在模型中的单词进行相关性的计算，并把相关性又编码叠加在每个单词中。这样，大模型能够更好地理解和生成自然文本，同时还能够表现出一定的逻辑思维和推理能力



大模型

基于深度学习

利用

大量的数据和计算资源

训练

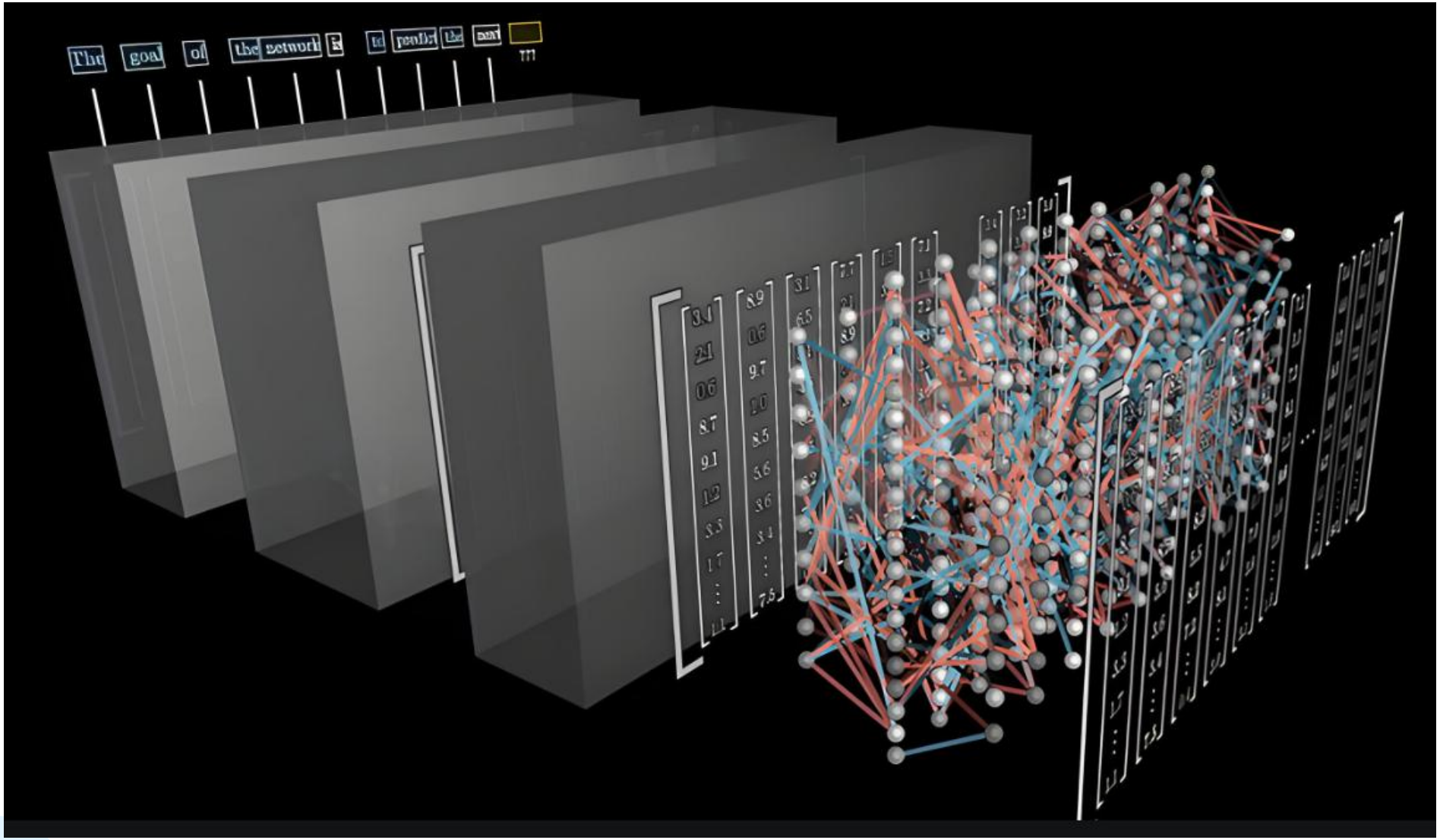
具有大量参数的
神经网络模型

不断地调整

模型参数

模型能够在各种任务中取得最佳表现

3.5 大模型的基本原理

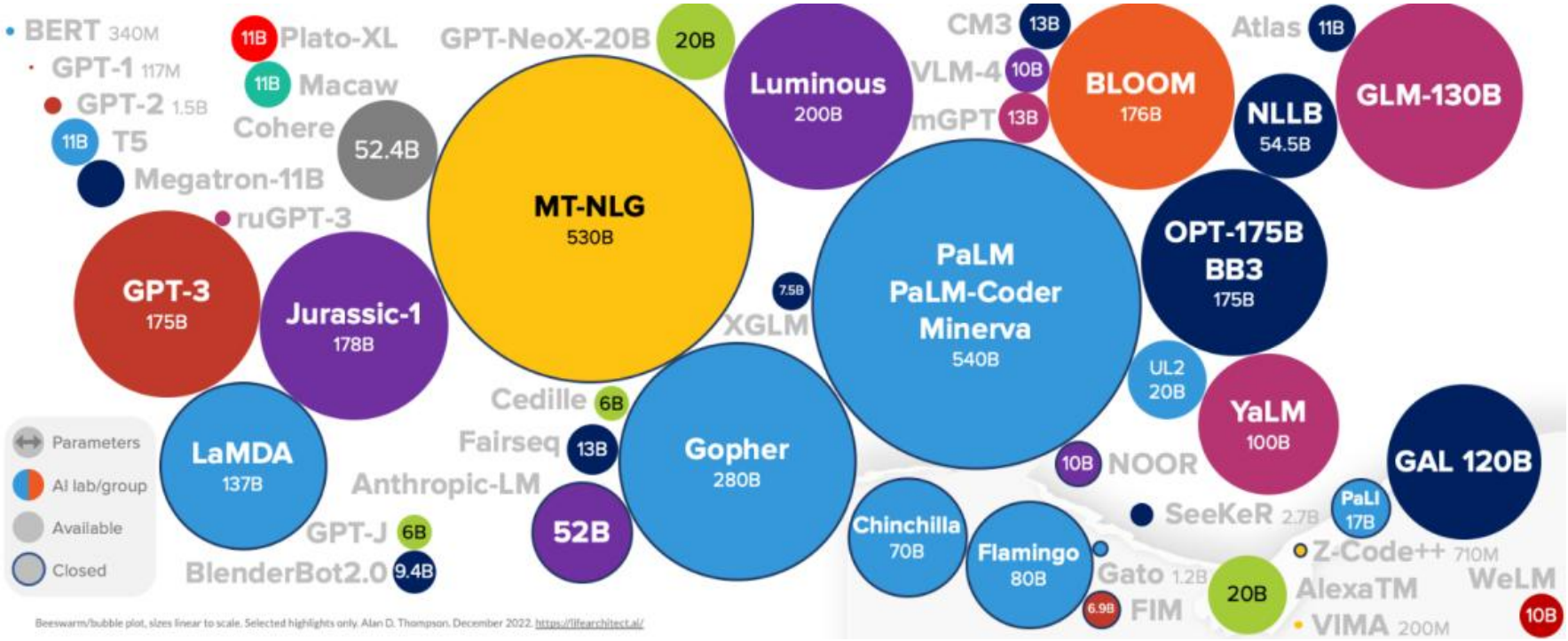


3.6 大模型的特点

3.6 大模型的特点

（1）巨大的规模

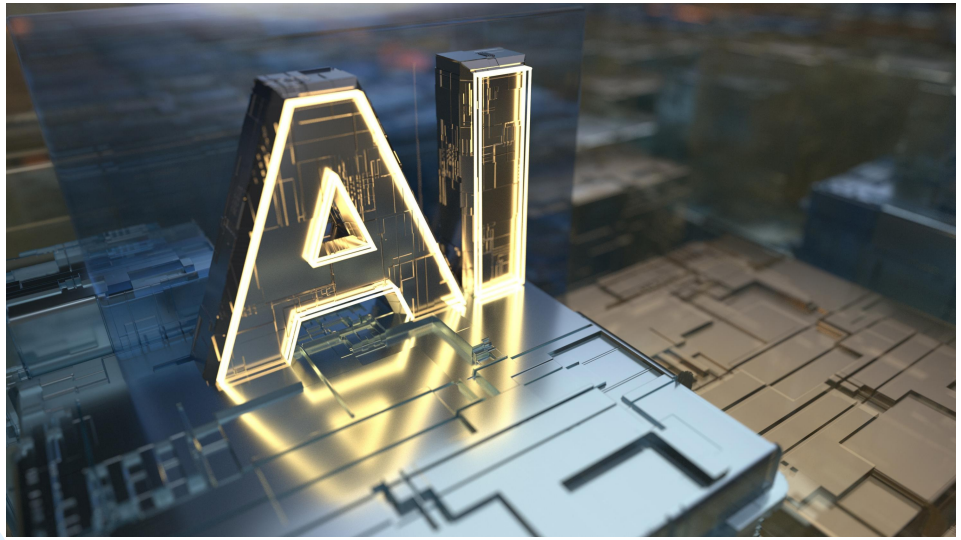
大模型通常包含数十亿个参数，模型大小可以达到数百GB甚至更大。这种巨大的规模不仅提供了强大的表达能力和学习能力，还使得大模型在处理复杂任务时具有更高的效率和准确性

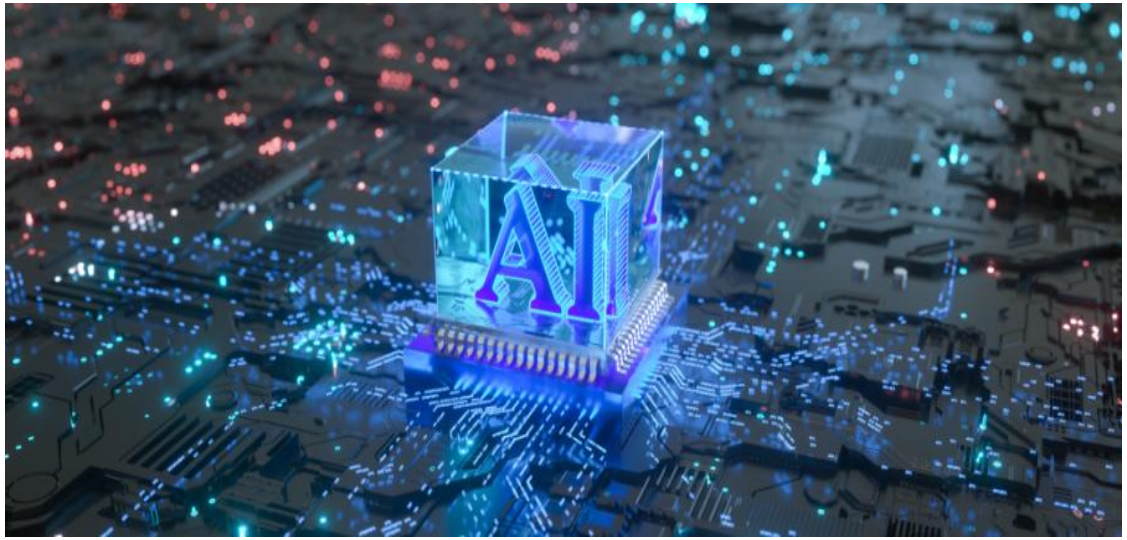


3.6 大模型的特点

（2）涌现能力

涌现能力是指模型在训练过程中突然展现出之前小模型所没有的、更深层次的复杂特性和能力。当模型的训练数据突破一定规模时，模型能够综合分析和解决更深层次的问题，展现出类似人类的思维和智能。**这种涌现能力是大模型最显著的特点之一，也是其超越传统模型的关键所在**





3.6 大模型的特点

（3）更好的性能和泛化能力

大模型因其巨大的规模和复杂的结构，展现出更出色的性能和泛化能力。它们在各种任务上都能表现出色，超越了传统的小模型。**这主要归功于大模型的参数规模和学习能力**。大模型能够更好地理解和模拟现实世界中的复杂现象，从而在各种任务中表现出更高的准确性和效率。它们能够捕捉到数据中的微妙差异和复杂模式，使得在未见过的数据上也能表现优秀，即具有良好的泛化能力

（4）多任务学习

大模型的多任务学习特点使其能够同时处理多种不同的任务，并从中学习到更广泛和泛化的语言理解能力。通过多任务学习，大模型可以在不同的NLP（Natural Language Processing）任务中进行训练，例如机器翻译、文本摘要、问答系统等。**这种多任务学习的方式有助于大模型更好地理解和应用语言的规则和模式**

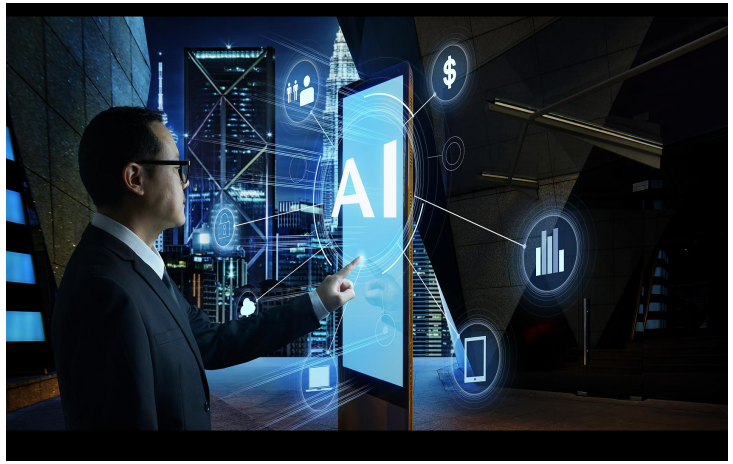
3.6 大模型的特点

（5）大数据训练

大模型需要大规模的数据来训练，通常在TB级别甚至PB级别。这是因为大模型拥有数亿甚至数十亿的参数，需要大量的数据来提供足够的信息供模型学习和优化。只有大规模的数据才能让大模型的参数规模发挥优势，提高模型的泛化能力和性能。同时，大数据训练也是保证大模型能够处理复杂任务的关键。通过使用大规模数据，大模型能够更好地理解数据中的复杂模式和关系，从而更好地模拟现实世界中的各种现象

（6）强大的计算资源

大模型需要强大的计算资源来训练和运行。由于模型规模庞大，参数数量众多，计算复杂度极高，因此需要高性能的硬件设备来支持。通常，训练大模型需要使用GPU（Graphics Processing Unit，图形处理器）或TPU（Tensor Processing Unit，张量处理器）等专用加速器来提高计算效率。这些加速器能够并行处理大量的参数和数据，使得大模型的训练和推断速度更快。除了硬件设备，大模型的训练还需要大量的时间。由于模型参数众多，训练过程中需要进行大量的迭代和优化，因此，训练周期可能长达数周甚至数月



3.6 大模型的特点

（7）迁移学习和预训练

通过在大规模数据上进行预训练，大模型能够学习到丰富的语言知识和模式，从而在各种任务上展现出卓越的性能。迁移学习和预训练有助于大模型更好地适应特定任务

在特定任务的数据上进行微调

在大规模数据上进行预训练后，大模型可以在特定任务的数据上进行微调，从而更好地适应目标任务的特性和要求。这种微调过程可以帮助大模型更好地理解和处理目标任务的特定问题，进一步提高模型的性能

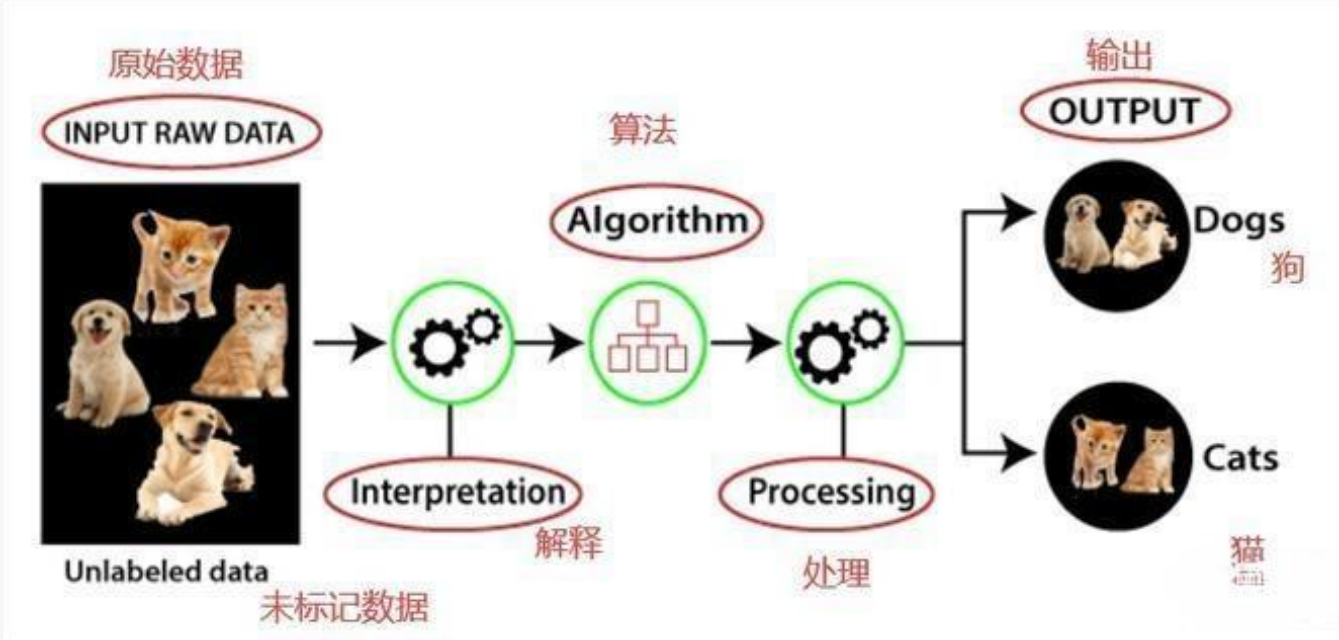
跨领域的应用能力

迁移学习和预训练也有助于大模型实现跨领域的应用。通过在多个领域的数据上进行预训练，大模型可以学习到不同领域的知识和模式，并在不同领域中进行应用。这种跨领域的应用能力，有助于大模型更好地服务于实际需求，推动人工智能技术的创新和发展

3.6 大模型的特点

（8）自监督学习

自监督学习利用大规模未标记数据进行训练，通过从数据中挖掘内在的规律和模式，使模型能够自动地理解和预测数据中的信息。在大规模的未标记数据中，大模型通过预测输入数据的标签或下一个时刻的状态来进行训练。这种训练方式使得大模型能够从大量的数据中自动地学习到语言的内在结构和模式，而不需要人工标注和干预

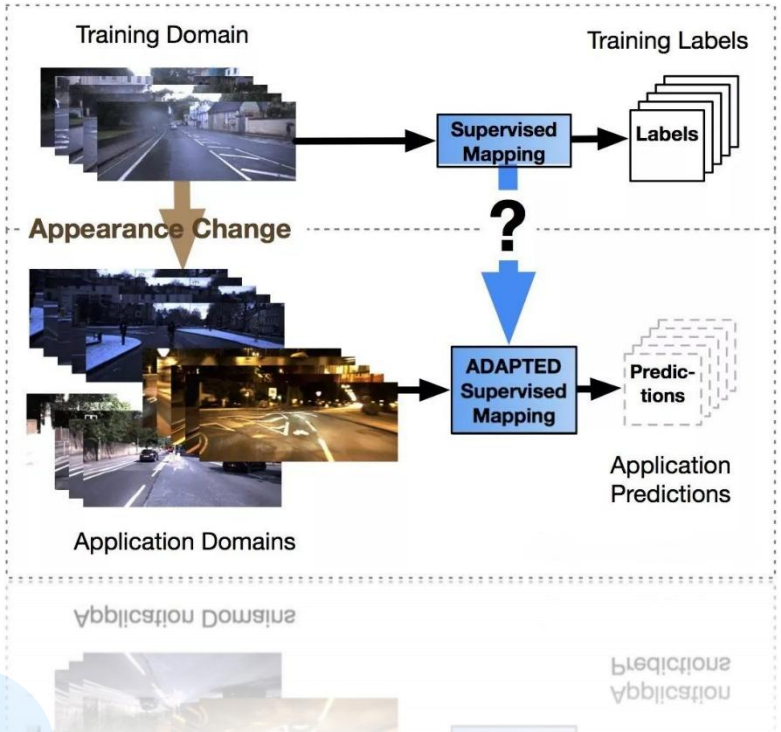


3.6 大模型的特点

（9）领域知识融合

大模型通过领域知识融合，能够将不同领域的数据和知识融合在一起，从而更好地模拟现实世界中的复杂现象

领域知识融合使得大模型能够从多个领域中学习到广泛的知识和模式，并将这些知识和模式整合到统一的框架中

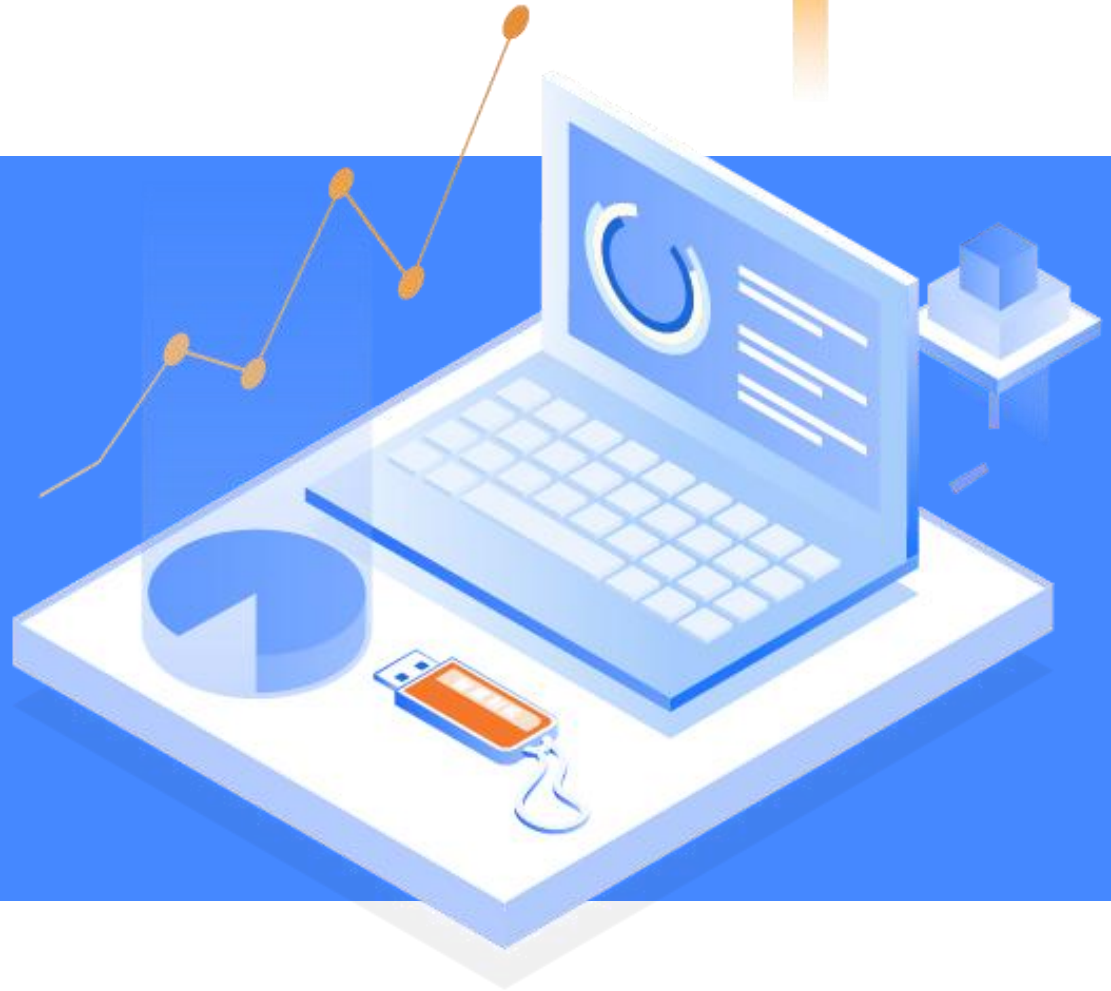


3.6 大模型的特点

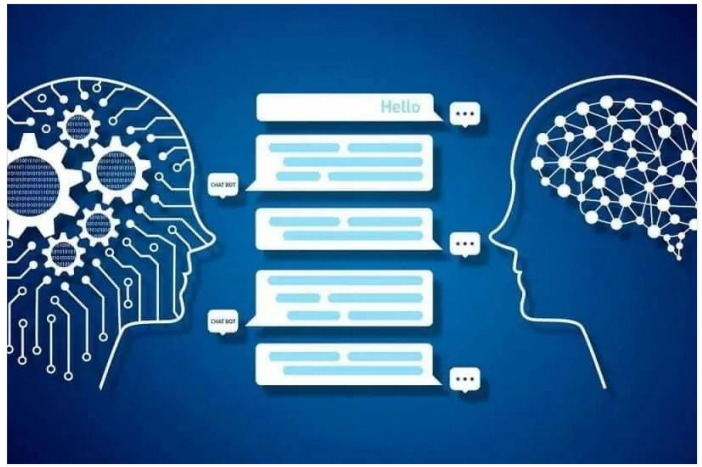
（10）自动化和效率

大模型在应用中展现出高度的自动化和效率。由于大模型具有强大的表达能力和学习能力，它可以自动化许多复杂的任务，大大提高工作效率。大模型通过预训练和微调过程，能够自动地适应特定任务，而不需要过多的手工调整和干预。这使得大模型能够快速地应用于各种实际场景，并且自动地处理复杂的任务，如自动编程、自动翻译、自动摘要等

3.7 大模型的分类



3.7 大模型的分类

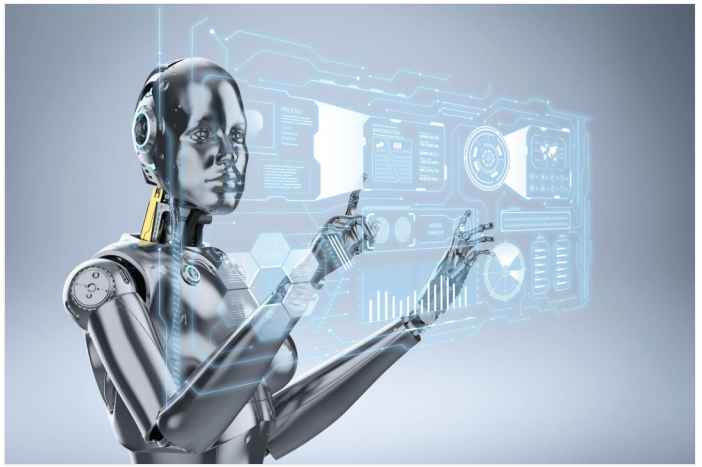


语言大模型

是指在自然语言处理（Natural Language Processing，NLP）领域中的一类大模型，通常用于处理文本数据和理解自然语言。这类大模型的主要特点是它们在大规模语料库上进行了训练，以学习自然语言的各种语法、语义和语境规则。代表性产品包括GPT系列（OpenAI）、Bard（Google）、DeepSeek、文心一言（百度）等

视觉大模型

是指在计算机视觉（Computer Vision，CV）领域中使用的大模型，通常用于图像处理和分析。这类模型通过在大规模图像数据上进行训练，可以实现各种视觉任务，如图像分类、目标检测、图像分割、姿态估计、人脸识别等。代表性产品包括VIT系列（Google）、文心UFO、华为盘古CV、INTERN（商汤）等



多模态大模型

是指能够处理多种不同类型数据的大模型，例如文本、图像、音频等多模态数据。这类模型结合了NLP和CV的能力，以实现对多模态信息的综合理解和分析，从而能够更全面地理解和处理复杂的数据。代表性产品包括DingoDB多模向量数据库（九章云极DataCanvas）、DALL-E(OpenAI)、悟空画画（华为）、midjourney等

3.7 大模型的分类

按照应用领域的不同，大模型主要可以分为L0、L1、L2三个层级

通用大模型L0

是指可以在多个领域和任务上通用的大模型。它们利用大算力、使用海量的开放数据与具有巨量参数的深度学习算法，在大规模无标注数据上进行训练，以寻找特征并发现规律，进而形成可“举一反三”的强大泛化能力，可在不进行微调或少量微调的情况下完成多场景任务，相当于AI完成了“通识教育”

行业大模型L1

是指那些针对特定行业或领域的大模型。它们通常使用行业相关的数据进行预训练或微调，以提高在该领域的性能和准确度，相当于AI成为“行业专家”

垂直大模型L2

是指那些针对特定任务或场景的大模型。它们通常使用任务相关的数据进行预训练或微调，以提高在该任务上的性能和效果

3.8 大模型的应用领域

3.8 大模型的应用领域

大模型的应用领域非常广泛，涵盖了自然语言处理、计算机视觉、语音识别、推荐系统、医疗健康、金融风控、工业制造、生物信息学、自动驾驶、气候研究等多个领域

（1）自然语言处理

大模型在自然语言处理领域具有重要的应用，可以用于文本生成（如文章、小说、新闻等的创作）、翻译系统（能够实现高质量的跨语言翻译）、问答系统（能够回答用户提出的问题）、情感分析（用于判断文本中的情感倾向）、语言生成（如聊天机器人）等

（2）计算机视觉

大模型在计算机视觉领域也有广泛应用，可以用于图像分类（识别图像中的物体和场景）、目标检测（能够定位并识别图像中的特定物体）、图像生成（如风格迁移、图像超分辨率增强）、人脸识别（用于安全验证和身份识别）、医学影像分析（辅助医生诊断疾病）等

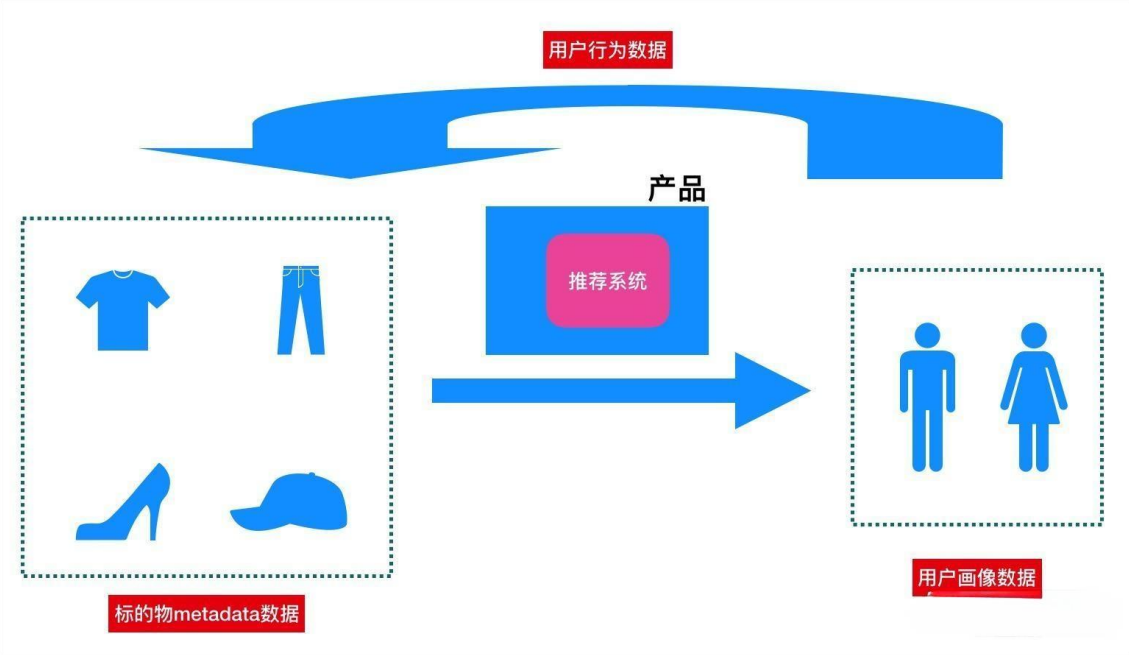
3.8 大模型的应用领域

（3）语音识别

大模型在语音识别领域也有应用，如语音识别、语音合成等。通过学习大量的语音数据，大模型可以实现高质量的跨语言翻译和语音识别以及生成自然语音

（4）推荐系统

大模型可以用于个性化推荐、广告推荐等任务。通过分析用户的历史行为和兴趣偏好，大模型可以为用户提供个性化的推荐服务，提高用户满意度和转化率



3.8 大模型的应用领域

自动驾驶

大模型可以用于自动驾驶中的感知、决策等任务。通过学习大量的驾驶数据，大模型可以实现对车辆周围环境的感知和识别，以及进行决策和控制，提高自动驾驶的安全性和效率

医疗健康

大模型可以用于医疗影像诊断、疾病预测等任务。通过学习大量的医学影像数据，大模型可以辅助医生进行疾病诊断和治疗方案制定，提高医疗水平和效率

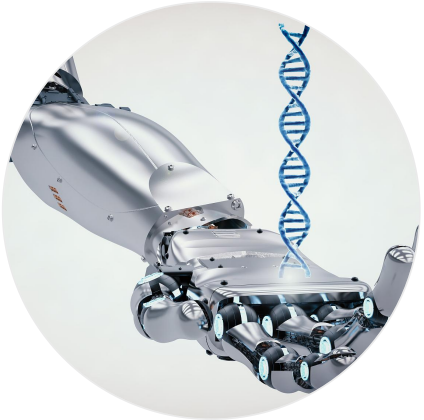
金融风控

大模型可以用于信用评估、欺诈检测等任务。通过分析大量的金融数据，大模型可以评估用户的信用等级和风险水平，以及检测欺诈行为，提高金融系统的安全性和稳定性

3.8 大模型的应用领域

工业制造

大模型可以用于质量控制、故障诊断等任务。通过学习大量的工业制造数据，大模型可以辅助工程师进行产品质量控制和故障诊断，提高生产效率和产品质量



生物信息学

在生物信息学领域，大模型可以用于基因序列分析（识别基因中的功能元件和变异位点）、蛋白质结构预测（推测蛋白质的二级和三级结构）、药物研发（预测分子与靶点的相互作用）等

气候研究

在气候研究领域，大模型可以处理气象数据，进行天气预测和气候模拟。它们能够分析复杂的气象现象，提供准确的气象预报，帮助人们做出应对气候变化的决策

3.9大模型对人们工作和生活的影响

3.9.1大模型对工作的影响
3.9.2 大模型对生活的影响

3.9.1大模型对工作的影响

提高工作效率

大模型在自然语言处理、机器翻译等领域的应用，使得人们能够快速、准确地处理大量文本数据，提高工作效率。例如，在翻译领域，大模型能够自动翻译多种语言，减少人工翻译的时间和成本，提高翻译效率

优化决策过程

大模型能够收集、整理和分析大量的数据，通过数据挖掘和机器学习技术，帮助人们更准确地了解问题现状，预测未来趋势，从而做出更明智的决策

自动化部分工作

大模型的发展使得一些繁琐、重复的工作可以由机器来完成，从而减轻了人们的工作负担。例如，在金融领域，大模型可以自动分析大量的金融数据，帮助人们做出更准确的决策

创造新的就业机会

随着大模型的普及和应用，将创造出许多新的就业机会。例如，需要更多的人来开发和维护大模型，也需要更多的人来利用大模型进行各种应用开发

3.9.2 大模型对生活的影响

改善生活质量

大模型在智能家居、智能客服等领域的应用，使得人们的生活更加便利、舒适。例如，通过智能家居系统，人们可以通过语音指令控制家电，实现智能化生活

提高学习效率

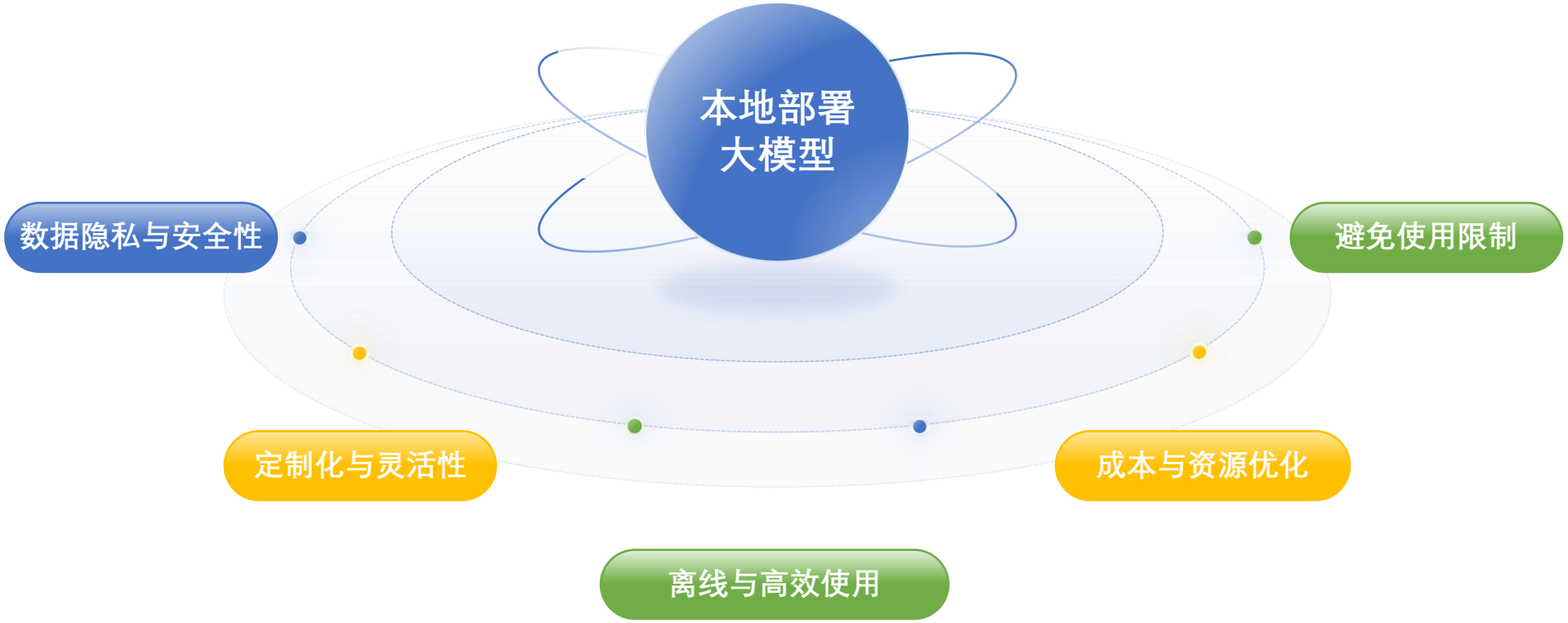
大模型在教育领域的应用，可以帮助人们更高效地学习新知识。例如，通过大模型的智能推荐功能，人们可以根据自己的兴趣和需求，获取更加个性化的学习资源

增强娱乐体验

大模型在娱乐领域的应用，可以提供更加丰富、多样的娱乐体验。例如，通过大模型的语音识别功能，人们可以通过语音指令控制游戏，实现更加智能化的游戏体验

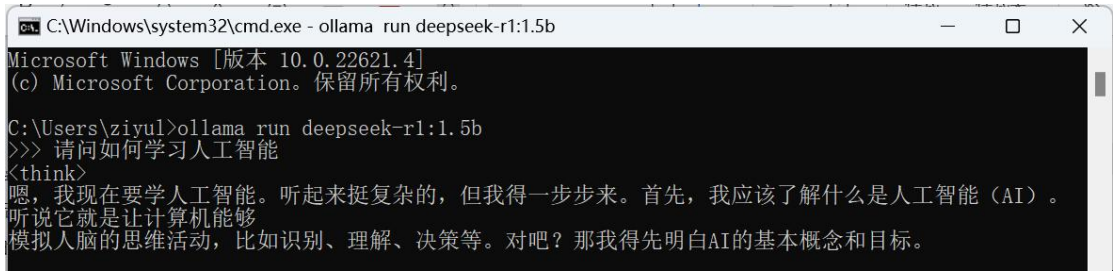
3.10 本地部署大模型

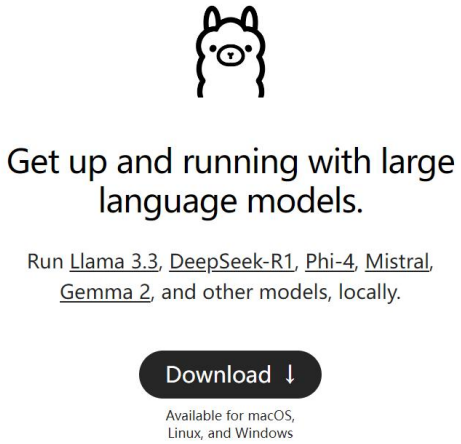
3.10.1 为什么需要本地部署大模型



3.10.2 本地部署大模型方法

- 安装Ollama
- 下载DeepSeek R1
- 运行DeepSeek R1
- 使用Open WebUI增强交互体验





具体安装过程请参考厦门大学数据库实验室博客 https://dblab.xmu.edu.cn/blog/5816/

3.11 基于大模型的智能体

3.11 基于大模型的智能体

智能体（AI Agent），又称“人工智能代理”，是一种模仿人类智能行为的智能化系统，它就像是拥有丰富经验和知识的“智慧大脑”，能够感知所处的环境，并依据感知结果，自主地进行规划、决策，进而采取行动以达成特定目标。简单来说，智能体能够根据外部输入做出决策，并通过与环境的互动，不断优化自身行为。

智能体本身既不是单纯的软件也不是硬件，而是一个更为宽泛的概念，它们可以是软件程序、机器人或其他形式的系统，具备一定的自主性和智能性。

3.11 基于大模型的智能体

基于大模型的智能体是指利用大语言模型（如GPT、BERT等）作为核心组件，构建的能够执行特定任务、与环境交互并做出决策的人工智能系统。这些智能体具有自主性、交互性、适应性等特点，能够模拟人类的认知和决策过程，提供更加自然、高效和个性化的交互体验。它们能够处理海量数据，进行高效的学习与推理，并展现出跨领域的应用潜力。

3.11 基于大模型的智能体

2025年1月23日，OpenAI发布了一个创新性的智能体——Operator，它是一个能够像人类一样使用计算机的智能体。它基于OpenAI最新研发的CUA（Computer-Using Agent）模型，CUA 将 GPT-4o 的视觉功能与通过强化学习获得的高级推理相结合，经过训练可以与图形用户界面（GUI，即人们在屏幕上看到的按钮、菜单和文本字段）进行交互。Operator通过观察屏幕并使用虚拟鼠标和键盘来完成任务，而无需依赖专门的API接口。这种设计使其可以适配任何为人类设计的软件界面，带来极高的灵活性。

Operator好比一个博士水平的个人助理，你给他一个复杂的任务，它就会自动执行。Operator的主要功能包括自主完成诸如采购杂货、提交费用报表、订票、买日用品、填写表格等任务，旨在通过自动化操作提升日常生活和工作效率。它还可以一边在StubHub搜索勇士队比赛门票，一边处理网球场预订、寻找清洁服务和DoorDash订餐，实现多任务并行处理。

3.11 基于大模型的智能体

2025年2月3日，OpenAI发布了一款新的智能体产品——Deep Research。Deep Research由OpenAI o3模型的一个版本提供支持，该模型针对网页浏览和数据分析进行了优化，它利用推理来搜索、解释和分析互联网上的大量文本、图像和PDF，并根据需要根据遇到的信息做出调整。Deep Research具有以下四大核心技术：

（1）数据雷达。会自动24小时扫描全球知识库。

（2）知识拼图。能把零散的信息拼成完整的战略地图。

（3）逻辑推理。发现矛盾时，自动回溯、验证，调整推理路径。

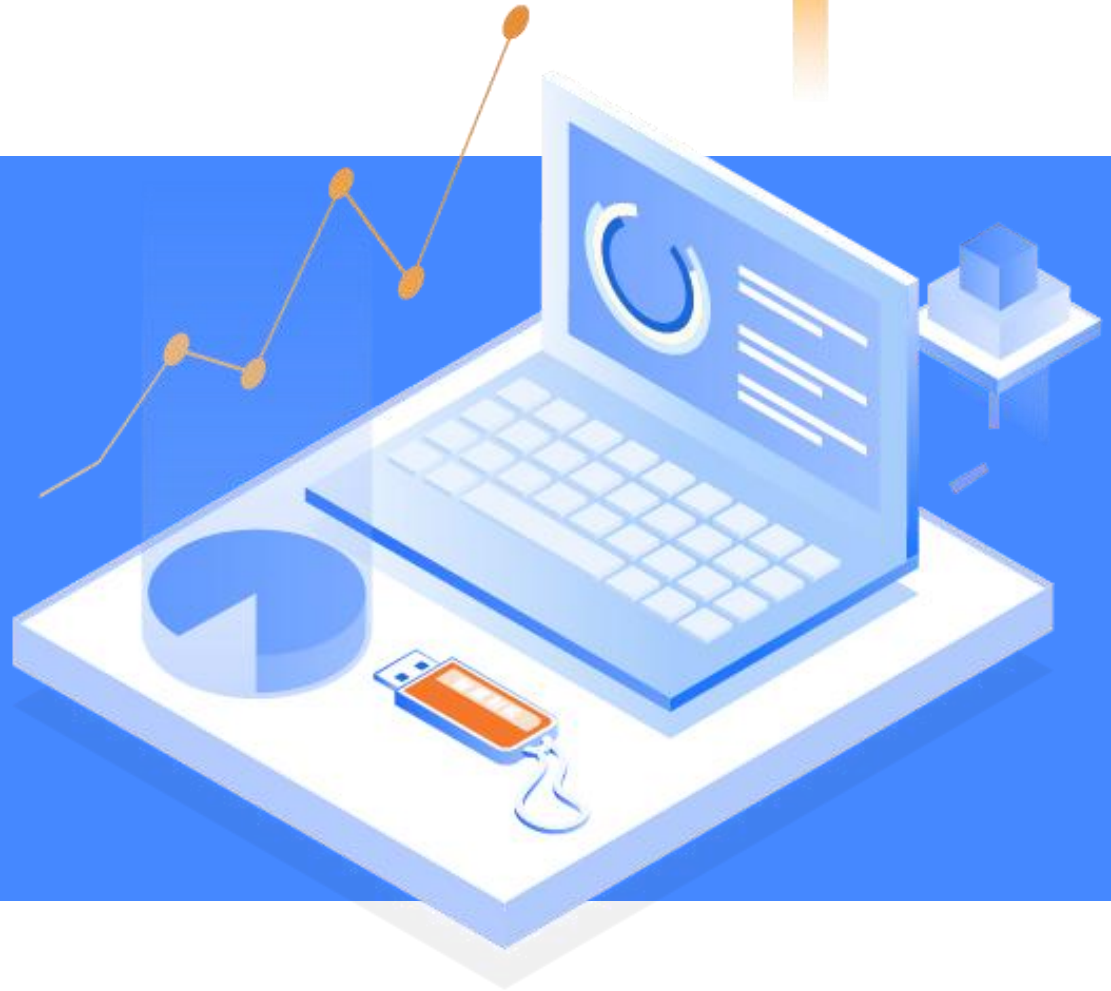
（4）学术裁缝。可以综合各种知识，生成完美的报告，还附带文献引用。

3.11 基于大模型的智能体

Coze(扣子)

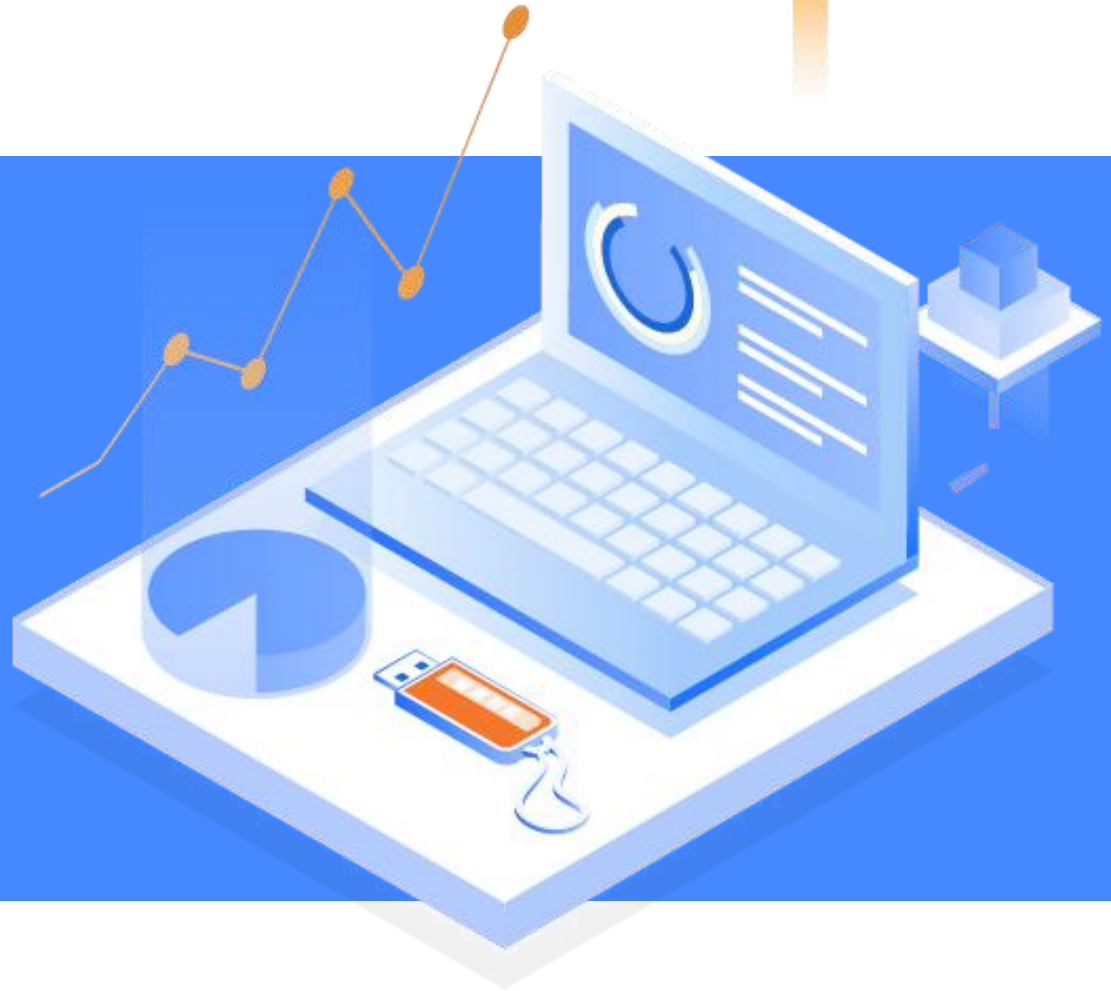
4. AIGC应用与实践

4.1 AIGC概述
4.2 文本类AIGC应用实践
4.3 图片类AIGC应用实践
4.4 语音类AIGC应用实践
4.5 视频类AIGC应用实践
4.6 AIGC在辅助编程中的应用
4.7 AI搜索
4.8 AI智能办公



4. 1 AIGC概述

4.1.1 什么是AIGC
4.1.2 AIGC与大模型的关系
4.1.3 常见的AIGC应用场景
4.1.4 AIGC技术对行业发展的影响
4.1.5 AIGC技术对职业发展的影响
4.1.6 常见的AIGC大模型工具
4.1.7 AIGC大模型的提示词



4.1.1 什么是AIGC

AIGC的全称为“Artificial Intelligence Generated Content”，中文翻译为“人工智能生成内容”。这是一种新的创作方式，利用人工智能技术来生成各种形式的内容，包括文字、音乐、图像、视频等

- AIGC是**人工智能进入全新发展时期的重要标志**，其核心技术包括生成对抗网络（GAN，Generative Adversarial Networks）、大型预训练模型、多模态技术等
- AIGC的核心思想是**利用人工智能算法生成具有一定创意和质量的内容**。通过训练模型和大量数据的学习，AIGC可以根据输入的条件或指导，生成与之相关的内容。例如，通过输入关键词、描述或样本，AIGC可以生成与之相匹配的文章、图像、音频等
- AIGC技术**不仅可以提高内容生产的效率和质量，还可以为创作者提供更多的灵感和支持**。在文学创作、艺术设计、游戏开发等领域，AIGC可以自动创作出高质量的文本、图像和音频等内容。同时，AIGC也可以应用于媒体、教育、娱乐、营销、科研等领域，为用户提供高质量、高效率、高个性化的内容服务

4.1.2 AIGC与大模型的关系

大模型与AIGC之间的关系可以说是**相辅相成、相互促进**的。大模型为AIGC提供了强大的技术基础和支撑，而AIGC则进一步推动了大模型的发展和应用，具体如下：

01

大模型为AIGC提供了丰富的数据资源和强大的计算能力

02

AIGC的需求也推动了大模型的发展

03

大模型和AIGC的结合，也带来了广泛的应用前景

4.1.3 常见的AIGC应用场景

AIGC可以应用于各行各业，主要包括但不限于生成文字、图像、音频、视频等，具体如下：

电商

生成商品标题、描述、广告文案和广告图

办公

写周报日报，写方案，写运营活动，制作PPT，写读后感，写代码

游戏

生成场景原画，生成角色形象，生成世界观，生成数值，生成3D模型，生成NPC对话，音效生成

娱乐

头像生成，照片修复，图像生成，音乐生成

影视

生成分镜头脚本，生成剧本脚本，台词润色，生成推广宣传物料，音乐生成

4.1.3 常见的AIGC应用场景

动漫

原画绘制，动画生成，分镜生成，音乐生成

艺术

写诗，写小说，生成艺术创作品，草图生成，艺术风格转换，音乐创作

教育

批改试卷，试卷创建，搜题答题，课程设计，课程总结，虚拟讲师

设计

UI设计，美术设计，插画设计，建筑设计

媒体

软文撰写，大纲提炼，热点撰写

生活

制定学习计划，做旅游规划

4.1.4 AIGC技术对行业发展的影响

AIGC技术对行业发展的影响深远且广泛，主要体现在以下几个方面：

内容创作领域的革新

AIGC技术能够自动生成高质量的文本、图像、音频和视频等内容，极大地提高了内容创作的效率。在新闻、广告、自媒体等领域，AIGC已经实现了广泛应用，帮助创作者快速生成多样化、个性化的内容，满足市场需求。这种技术革新不仅降低了内容创作的成本，还激发了创作者的创新灵感，推动了内容产业的繁荣发展。

生产力提升与成本降低

AIGC技术在多个行业中展现了其提升生产力和降低成本的潜力。例如，在游戏开发领域，AIGC技术可以用于场景构建、角色互动等，减少人工制作的工作量，提高开发效率。在制造业中，AIGC技术可以辅助设计、优化生产流程，降低生产成本。这些应用使得企业能够更快地响应市场变化，提升竞争力。

用户体验的升级

AIGC技术通过提供个性化、定制化的内容和服务，显著提升了用户体验。在智能客服、在线教育等领域，AIGC技术可以根据用户的需求和偏好提供精准的服务，满足用户的个性化需求。这种以用户为中心的服务模式不仅增强了用户的满意度和忠诚度，还为企业带来了更多的商业机会。

推动行业创新与转型

AIGC技术的快速发展为传统行业带来了转型升级的契机。通过与AIGC技术的深度融合，传统行业可以探索新的商业模式和服务模式，实现创新发展。例如，在零售业中，AIGC技术可以用于智能推荐、虚拟试衣等场景，提升购物体验并促进销售增长。在金融领域，AIGC技术可以应用于投资策略优化、风险管理等方面，提高金融机构的决策效率和准确性。

4.1.5 AIGC技术对职业发展的影响

AIGC技术对职业发展产生了深远的影响，主要体现在以下几个方面：

新兴职业的出现

随着AIGC技术的快速发展，一系列与该技术相关的新兴职业应运而生。例如，AI训练师、机器学习工程师、数据标注员等职业需求激增。这些新兴职业不仅要求从业者具备扎实的技术基础，还需要不断学习和掌握最新的AIGC技术动态。

传统职业的转型升级

AIGC技术也为传统职业的转型升级提供了契机。许多传统职业如编辑、设计师、教师等，在AIGC技术的辅助下，工作效率和创作质量得到了显著提升。同时，这些职业也需要从业者不断适应技术变革，掌握新的技能和工具，以适应市场需求的变化。

工作方式的变革

AIGC技术改变了传统的工作方式，使得远程工作、灵活办公成为可能。许多企业开始采用AIGC技术来优化工作流程，减少人力成本，提高工作效率。这种变革不仅为员工提供了更加灵活的工作方式，也为企业带来了更大的经济效益。

职业发展路径的多样化

AIGC技术的发展为职业发展路径提供了更多的可能性。从业者可以根据自己的兴趣和特长，选择适合自己的职业发展方向。例如，一些对AI技术感兴趣的从业者可以选择成为AI训练师或机器学习工程师，而一些具有创意和设计才能的从业者则可以利用AIGC技术来提升自己的创作能力。

持续学习与技能提升

面对AIGC技术的快速发展，从业者需要不断学习和提升自己的技能水平。通过参加培训课程、阅读专业书籍、参与技术论坛等方式，从业者可以紧跟技术前沿，保持自己的竞争力。

4.1.6 常见的AIGC大模型工具

常见的AIGC大模型工具包括：

- OpenAI的ChatGPT
- DeepSeek
- 科大讯飞的讯飞星火
- 阿里的通义千问
- 百度文心一言
- 字节跳动豆包
- Kimi

这些工具基于大规模语言模型技术，**具备文本生成、语言理解、知识问答、逻辑推理等多种能力，可广泛应用于写作辅助、内容创作、智能客服等多个领域**。通过不断迭代和优化，为用户提供更加智能、高效的内容生成解决方案

4.1.7 AIGC大模型的提示词

- AIGC大模型的提示词（Prompt）是指**用户向大模型输入的文本内容，用于触发大模型的响应并指导其如何生成或回应**
- 这些提示词可以是一个问题、一段描述、一个指令，甚至是一个带有详细参数的文字描述。它们为大模型提供了生成对应文本、图片、音频、视频等内容的基础信息和指导方向。
- 提示词的重要作用如下：

引导生成

增强交互性

提高准确性

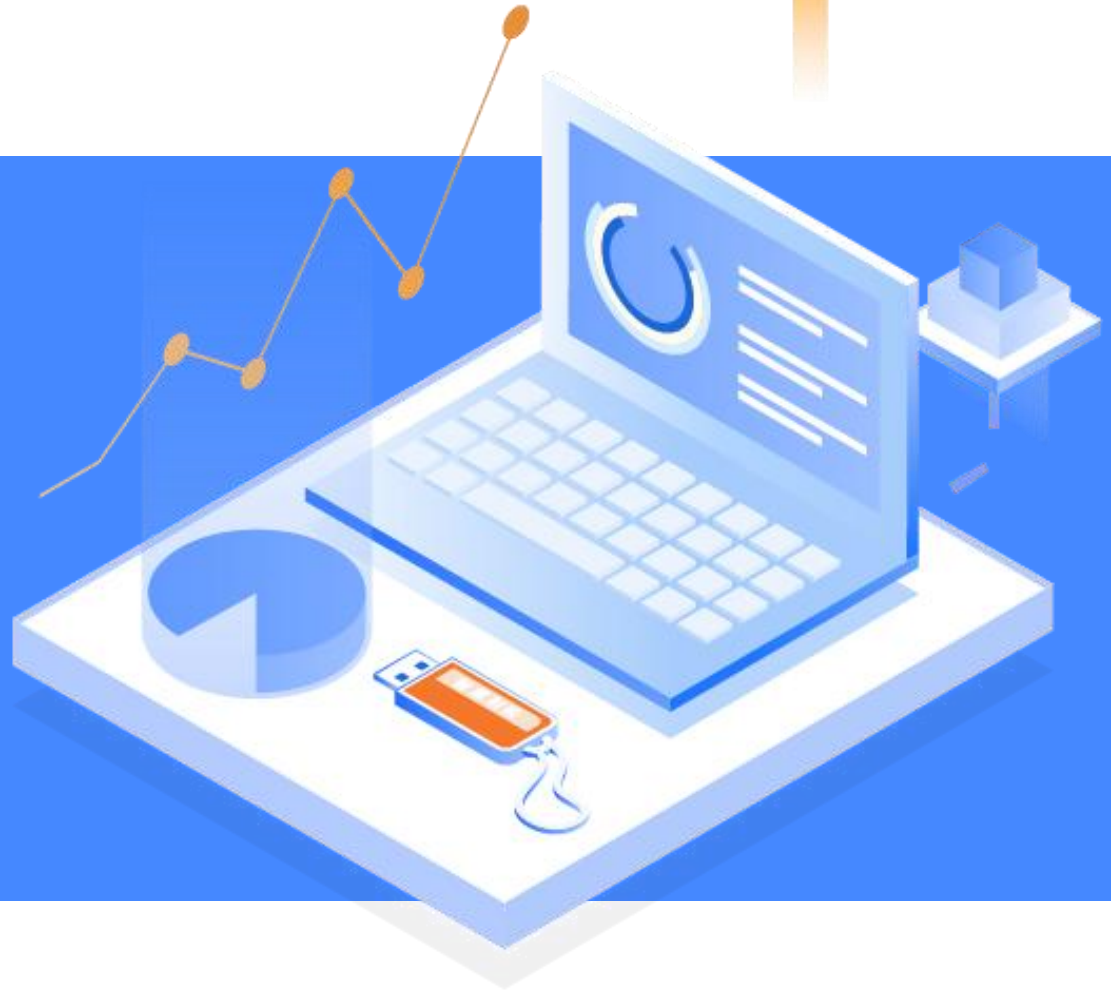
4.1.7 AIGC大模型的提示词

使用提示词需要注意一些技巧，这样可以从大模型获得更加符合我们预期要求的结果，主要技巧如下：

- 简洁明确
- 考虑受众
- 分解复杂任务
- 使用肯定性指令
- 示例驱动
- 明确角色
- 遵守规则
- 自然语言回答

4.2 文本类AIGC应用实践

4.2.1 案例1：与DeepSeek进行对话
4.2.2 案例2：与百度文心一言进行对话
4.2.3 案例3：使用讯飞智文生成PPT



4.2.1 案例1：与DeepSeek进行对话

1.快速体验DeepSeek

访问DeepSeek官网（https://chat.deepseek.com/），会出现如图7-1所示对话界面，在提示词输入框的底部，有两个按钮，即“深度思考(R1)”和“联网搜索”，可以用鼠标点击来选中或取消，默认情况下，“深度思考(R1)”按钮是处于选中状态，“联网搜索”则处于未选中状态。两个按钮的功能如下：

（1）深度思考(R1)。表示触发更复杂的多步推理能力，适合需要逻辑链分析的场景，典型使用场景包括数学题/物理题推导、文学作品的隐喻分析、编程问题的架构设计、需要分步骤解释的操作指南等。

（2）联网搜索。表示实时获取最新网络信息，适合时效性强的查询，典型使用场景包括查询实时股价/汇率、验证最新科研成果、获取突发事件进展、检索特定网页内容等。

4.2.1 案例1：与DeepSeek进行对话

1.快速体验DeepSeek

在提示词输入框中输入“请模仿李白的《望庐山瀑布》做一首诗，题目是《看厦门鼓浪屿》”，然后回车，或者用鼠标点击提示词输入框右侧的箭头按钮，向DeepSeek发起提问。DeepSeek给出的回答如图7-2所示，需要注意的是，大模型属于概率模型，每次生成的回答内容可能不完全相同。

已深度思考（用时 21 秒）

《看厦门鼓浪屿》

鹭岛东南烟水遥，日光岩上碧云飘。

琴潮忽卷千堆雪，散作凤凰霞万条。

注：本诗仿太白雄奇飘逸之风，首句以“烟水遥”勾勒海上仙山意境，承以日光岩实景作天地支点。转句化琴岛声波为视觉浪涌，取“千堆雪”暗合东坡赤壁豪情，结句凤凰木朱华迸天，将声、光、色熔铸为漫天红霞，较之银河落九天另开海上生霞之瑰丽新境。

4.2.1 案例1：与DeepSeek进行对话

2.DeepSeek的基本用法

（1）基本原则：简单直接，自然表达。①无需复杂结构。直接描述需求即可，无需添加“角色扮演”（如“假设你是专家”）或复杂指令（如“用学术语言分三点回答”）。比如，你可以直接向DeepSeek提问“什么是光合作用？”、“如何用Python写一个计算器程序？”，而不建议使用提示词“请以生物学教授的身份，用三个段落解释光合作用，每段不超过100字”。②多轮对话优化结果。如果首次回答不完整，可通过追问补充细节，无需一次性给出完美提示。比如，第一轮提问“写一首关于秋天的诗”，第二轮提问“加入一些悲伤的情绪”，第三轮提问“把‘落叶’换成比喻句”。

（2）不同场景的提问技巧（非必需，但可提升效率）。虽然简单提问即可满足大多数需求，但在复杂任务中，适当提供背景信息或明确需求会让结果更精准，具体技巧包括：①知识类问题。比如，基础提问是“量子力学的基本原理是什么？”，优化后的提问是“用通俗易懂的语言解释量子纠缠，适合高中生理解”，② 创作类任务（写作、编程等）。比如，基础提问是“写一个关于人工智能的科幻短篇故事”，优化后的提问是“写一个反乌托邦主题的科幻故事，主角是女性工程师，结局有反转”。③实用建议（学习、工作等）。比如，基础提问是“如何提高英语听力？”，优化后的提问是“我每天只有30分钟学习时间，有哪些高效的英语听力练习方法？”。④复杂任务（数据分析、代码调试）。比如，基础提问是“这段Python代码报错了，帮我看看问题”，优化后的提问是“我的代码目标是爬取网页数据，但遇到SSL证书错误。报错信息如下：[粘贴代码]”。

4.2.1 案例1：与DeepSeek进行对话

2.DeepSeek的基本用法

作为初学者，DeepSeek的一些“魔法”指令也很有用，比如，你可以输入“/步骤 如何用手机拍摄旅游照片”，DeepSeek返回的回答结果就会按照步骤详细给出拍摄旅游照片的说明，再比如，你可以输入“请解释量子计算，然后/简化”，它就会返回比较简明扼要的回答。

表 DeepSeek的“魔法”指令

指令	功能
/续写	当回答中断时自动继续生成
/简化	将复杂内容转换成大白话
/示例	要求展示实际案例（特别是写代码时）
/步骤	让AI分步骤指导操作流程
/检查	帮你发现文档中的错误

4.2.1 案例1：与DeepSeek进行对话

3.使用DeepSeek处理文档

点击DeepSeek界面中的「回形针」图标上传文件，支持的文件类型包括文本类（PDF、DOCX、TXT、Markdown）、数据类（CSV、XLSX）和图像类（JPG、PNG）。然后，就可以在对话框中输入提示词，比如，可以输入“总结这份年报的三个核心要点”、“提取合同中的责任条款制成表格”、“对比文档A和文档B的市场策略差异”、“从实验报告中整理所有温度数据”、“请识别图片中的文字”等。也可以使用一些指令来处理文档，如表7-2所示。

表 用于文档处理的DeepSeek指令

功能	指令模板	应用场景
内容摘要	/总结 [文件名] 生成500字摘要	快速把握长篇文档核心内容
问答提取	/问答 [文件名] 第三章提到的技术参数是?	精准定位特定信息
数据可视化	/可视化 [文件名] 将销售数据生成折线图	转化表格数据为图形分析
跨文档对比	/对比 文件A vs 文件B 的政策差异	合同/论文查重对比

还可以要求DeepSeek对回答结果进行结构化输出，比如，可以输入如下提示词：

/解析文件 年度报告.docx
输出要求:
1. 按"营收/利润/成本"分类
2. 用Markdown表格对比近三年数据
3. 关键增长点用✅标注

4.2.2 案例2：与百度文心一言进行对话

（1）告诉文心一言你要的风格

在输入提示词时，明确指定你希望生成的文本内容的风格。这样，文心一言在理解并处理你的请求时，会更有针对性地调整其生成内容的风格，以满足你的具体需求。比如，可以使用提示词：

> 请按照要求写一篇200字左右关于云计算的介绍。注意事项：文章的受众是中学生，需要通俗易懂，语言风格需要幽默、风趣一些

想要生成不同语气风格的文字，可以在问题描述中加入你想要的语气风格作为限定条件，提示文心一言按照你的要求去输出。

比如，如果你需要正式语气，可以在提示词中加入“请采用正式的词汇和语法结构，使内容显得庄重、严肃和专业”；如果你需要抒情语气，请在提示词中加入“请使用富有感情和表达感情的词汇，使内容产生共鸣和情绪共振”；如果你需要口语化语气，请在提示词中加入“请运用口语化的表达方式，例如俚语、俗语和口头禅，使内容更加轻松和亲切”。

4.2.2 案例2：与百度文心一言进行对话

（2）告诉文心一言你要的结构

在构建提示词时，应明确指定期望的输出结构。比如，如果是要求生成一篇文章，可以在提示词中明确指出“请按照引言-正文-结论的结构来撰写”。这样，文心一言在生成内容时，会遵循这一结构框架，使得输出更加条理清晰、逻辑严密。再比如，如果要撰写给上级领导的方案、报告、总结时，可以使用提示词：

> 请按照【现状/问题/解决方案，数据洞察/问题概览/调研方向，数据/亮点/问题/经验】这个结构撰写一份关于我国芯片行业的总结报告

4.2.2 案例2：与百度文心一言进行对话

（3）告诉文心一言你要的角色

在提示词中可以设定具体的角色或视角。例如，在要求创作故事时，可以明确指定“以一位勇敢探险家的视角讲述这段经历”。这样的提示能引导文心一言在生成内容时，从特定角色的角度出发，赋予文本独特的情感色彩和叙事风格。此技巧有助于增强生成内容的代入感和故事性，使内容更加丰富和引人入胜。

下面是一段提示词实例：

> 请你作为一个小红书文案撰写高手，为我生成一篇爆款小红书文案，要求：突出酒店的特色，包括海景房、豪华单间、最新装修、免费早餐、无线上网等

下面是另一段提示词实例：

> 我希望你能扮演记者的角色，按照我的要求撰写一份新闻调查，要求：调查油罐车不清洗直接运送食用油的事情，不要出现具体企业名称，要给出政府部门的处理态度

4.2.2 案例2：与百度文心一言进行对话

（4）告诉文心一言你的内容要求

可以通过详细具体的提示词明确表达内容要求。无论是希望生成的文章主题、关键词汇，还是期望涵盖的信息点、情感倾向，都应在提示词中清晰呈现。这样做能让文心一言更准确地理解用户需求，生成更符合期望的内容。

比如，可以通过如下提示词表达自己的内容要求：

> 在6G专利申请方面，中国已经遥遥领先。2021年的数据显示，中国的6G专利申请量占比高达40.3%，稳坐世界第一的宝座。
> 请把上面的数据更新到目前最新的数据

4.2.2 案例2：与百度文心一言进行对话

（4）告诉文心一言你的内容要求

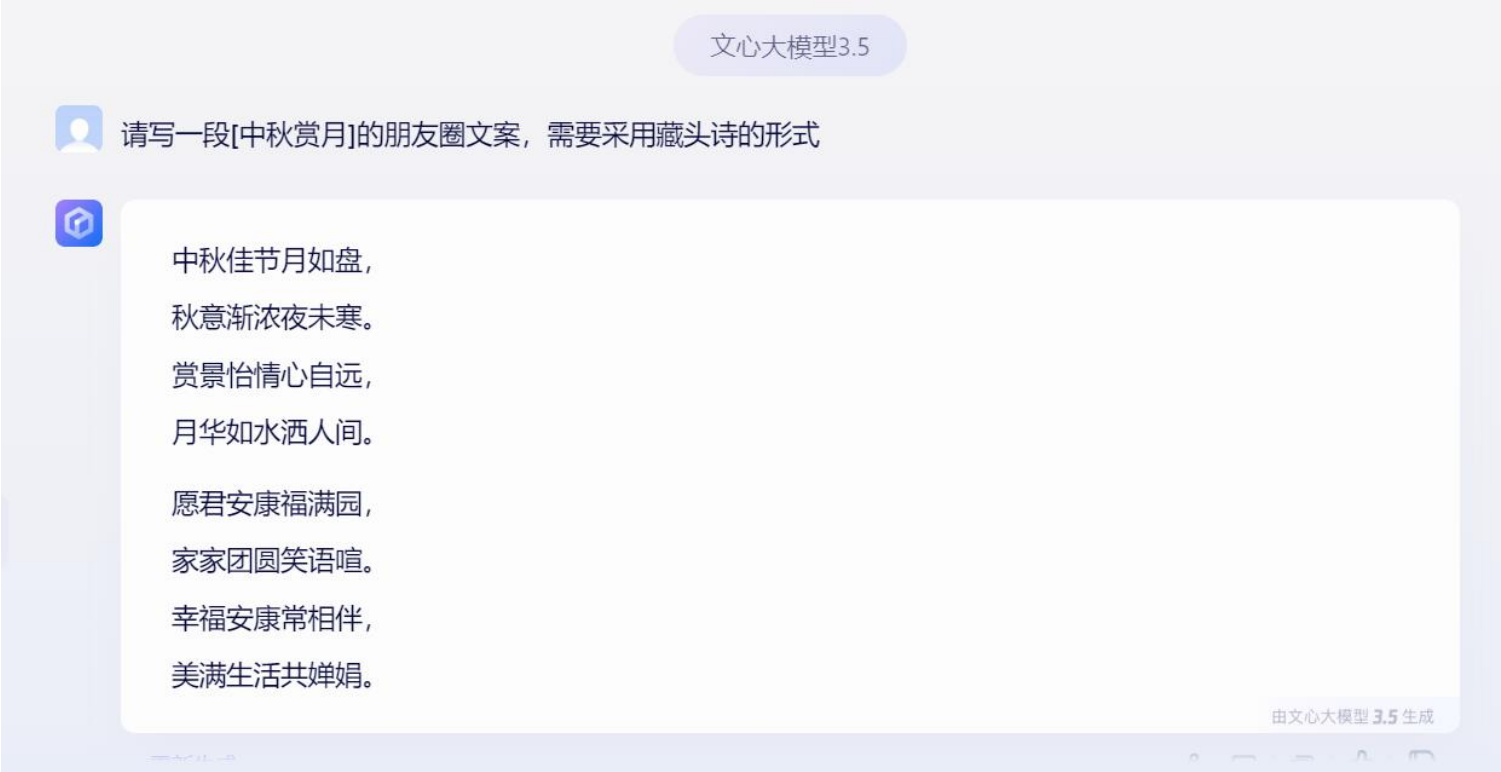
如果对输出的内容有比较多的要求或限制，不妨在输入框中将这些内容要求一条一条明确告诉文心一言，比如，可以采用类似如下的提示词：

请以小红书的风格，按照以下要求帮我为“海景美食餐厅”写一篇小红书种草文案；
内容要求：
（1）要有标题、正文
（2）标题字数：不超过20个字；尽量简短精炼，要足够吸引眼球，用词浮夸
（3）正文分段，层次分明，每段最少100字
（4）要用“首先、其次、最后”这种模式
（5）整篇文案不要超过1000个字

4.2.2 案例2：与百度文心一言进行对话

（5）告诉文心一言你想写的文体

明确指定文体，如散文、小说、诗歌、科技文等，让大模型理解并模拟该文体的语言特点、结构安排和表达习惯，从而输出更具针对性的文本。比如，可以采用提示词“请写一段[中秋赏月]的朋友圈文案，需要采用藏头诗的形式”



4.2.2 案例2：与百度文心一言进行对话

（6）指导文心一言分步解决问题

将复杂问题拆解成多个简单、具体的步骤，作为提示词输入给文心一言。这样不仅能降低问题的处理难度，使文心一言更容易理解和响应，还能确保解决问题的过程更加系统、有条理。通过逐步引导，可以逐步逼近问题的解决方案，提高答案的准确性和实用性。

比如，如果想让文心一言帮你制定一份旅行规划，可以使用类似如下的提示词：

请为我规划一次为期一周的厦门自由行；

（1）第1步：列出必去的景点，如厦门大学、鼓浪屿、环岛路、五缘湾、曾厝垵；

（2）第2步：根据景点位置安排每日行程，确保交通便利；

（3）第3步：推荐几家当地的特色餐厅，包括早餐、午餐和晚餐；

（4）第4步：提供一家性价比高的酒店住宿建议，并考虑其位置是否便于游览。

4.2.2 案例2：与百度文心一言进行对话

（7）告诉文心一言你要的示例

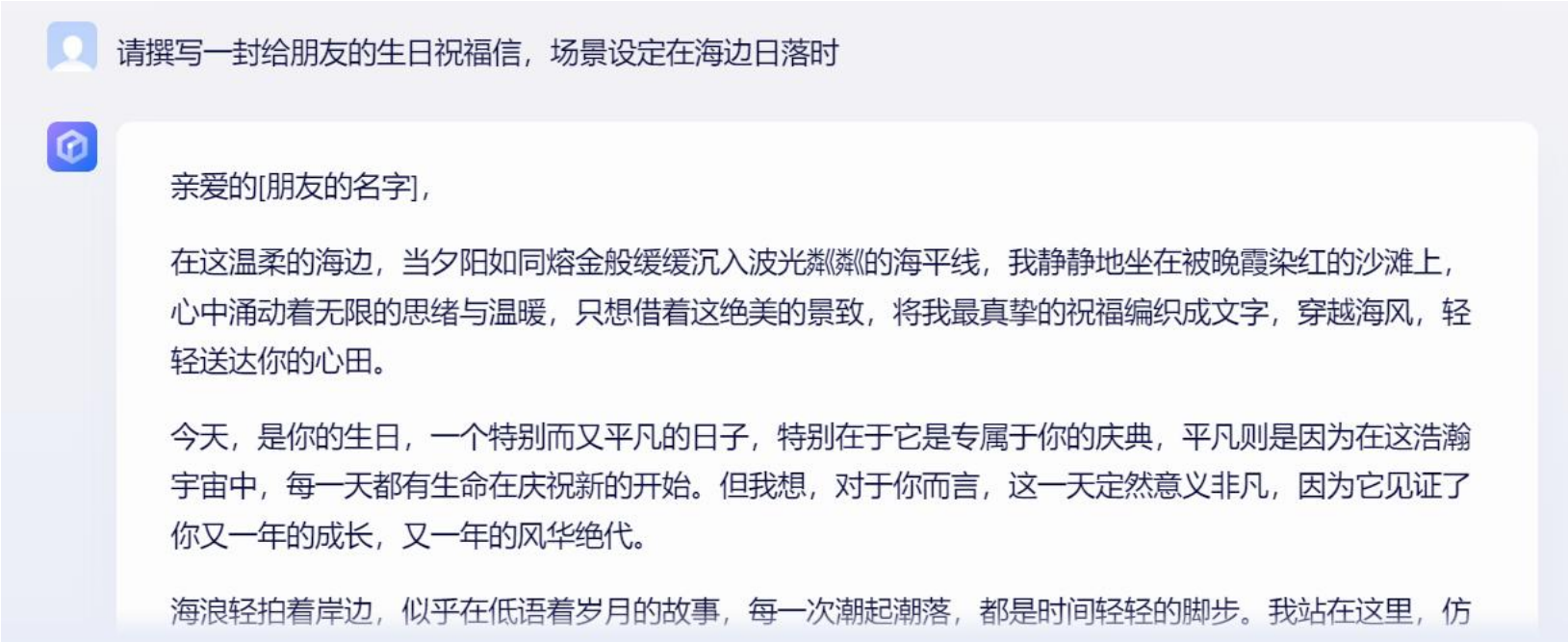
明确沟通意图，通过具体示例引导大模型理解你的需求。这有助于文心一言更准确地捕捉你的思维框架和期望结果，减少误解。比如，可以使用类似如下的提示词：

我是一位高校教师，请帮我写一份工作周报，内容尽量简洁精炼，下面是我本周的工作内容：
（1）完成了5个本科生毕业论文修改
（2）撰写了教材的一个章节“云计算与大数据”
输出要求示例：
【本周工作周报】
【本周工作进展】本周做了哪些事，产生了哪些结果
【下周工作安排】基于本周的结果下周要推进哪些事
【思考总结】简要说说本周的收获和反思

4.2.2 案例2：与百度文心一言进行对话

（8）告诉文心一言你要的场景

在输入提示词时，应明确描述所需的上下文或环境背景，如“在科幻电影中描述一个未来城市的景象”或“请撰写一封给朋友的生日祝福信，场景设定在海边日落时”。这样做有助于文心一言更好地理解你的需求，生成更符合场景氛围和情境的内容，从而提升输出内容的贴切性和情感共鸣。



4.2.3 案例3：使用讯飞智文生成PPT

讯飞智文是科大讯飞公司旗下的AI一键生成PPT/WORD的网站平台，是基于科大讯飞星火认知大模型技术基础上开发的一个具体应用，主要功能有文档一键生成、AI 撰写助手、多语种文档生成、AI自动配图、模板图示切换功能。这里介绍如何使用讯飞智文快速生成PPT。

请首先准备一个包含文本内容的PDF文件，比如，可以从网络新闻报道中复制一段关于2024年7月19日微软蓝屏事件的内容保存到一个WORD文档中，命名为“微软蓝屏.docx”，然后，使用WPS软件打开“微软蓝屏.docx”，把该WORD文档保存成PDF格式，生成“微软蓝屏.pdf”。

4.2.3 案例3：使用讯飞智文生成PPT

访问讯飞智文官网（https://zhiwen.xfyun.cn/），在首页（如图所示）中点击“免费使用”，然后按照网页提示完成注册（推荐使用手机号注册）。

4.2.3 案例3：使用讯飞智文生成PPT

在页面中（如图7-2 所示）点击“开始制作”。在出现的页面中（如图7-3所示），选择AI PPT的“文档创建”。然后，在出现的页面中（如图7-4所示），点击“点击上传”，把本地文件“微软蓝屏.pdf”文件上传上去（当然，也可以上传“微软蓝屏.docx”）。

然后，在出现的页面中（如图7-5所示），点击“开始解析文档”。之后，页面会显示提示文字“好的，已收到您的要求，让我先为您生成PPT标题和大纲”。过一会儿，就会显示自动生成的PPT标题和大纲，如果你不满意，可以点击页面底部的“重新生成”，如果满意，可以直接点击“下一步”。

图7-2 开始创作

4.2.3 案例3：使用讯飞智文生成PPT

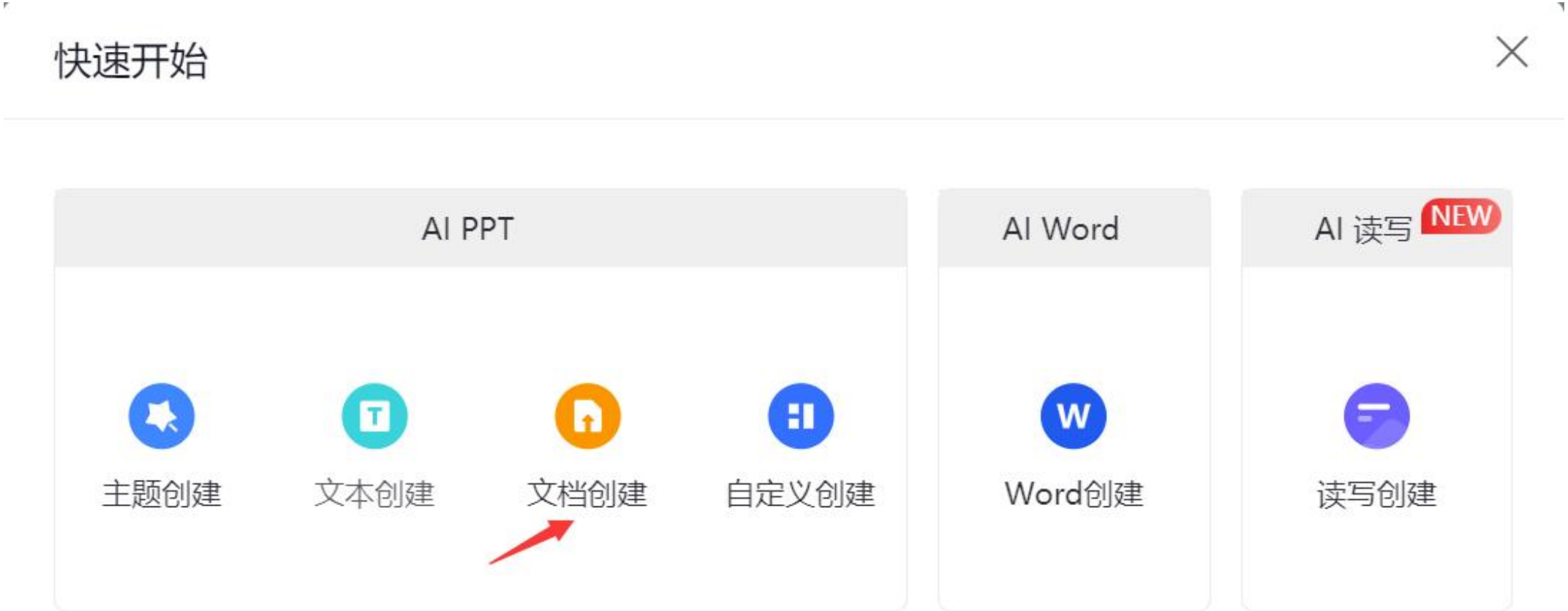


图7-3 选择AI PPT中的“文档创建”

4.2.3 案例3：使用讯飞智文生成PPT

图7-4 上传文件

4.2.3 案例3：使用讯飞智文生成PPT

图7-5 开始解析文档

4.2.3 案例3：使用讯飞智文生成PPT

在出现的页面中（如图7-6所示），选择你想要的模板配色，比如，这里选择“清逸天蓝”，然后点击页面顶部的“下一步”。经过一段时间以后，页面就会显示自动生成的PPT（如图7-7所示），点击页面右上角的“导出”，就可以把PPT保存到本地电脑中，然后，可以根据自己的需求，自己对PPT继续进行修改和完善。在本地电脑中打开自动生成的PPT，可以看出，AI制作PPT的水平非常专业，逻辑清晰，配图精美，超过了很多PPT初级者的制作水平，可以大大提高普通用户制作PPT的效率和水平。

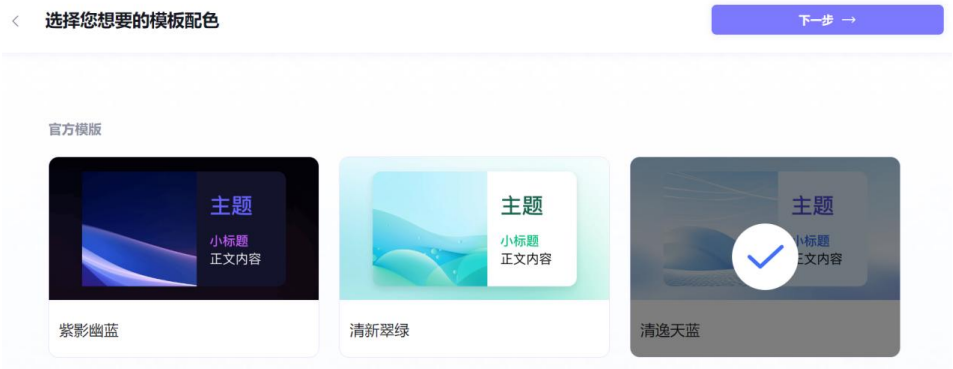
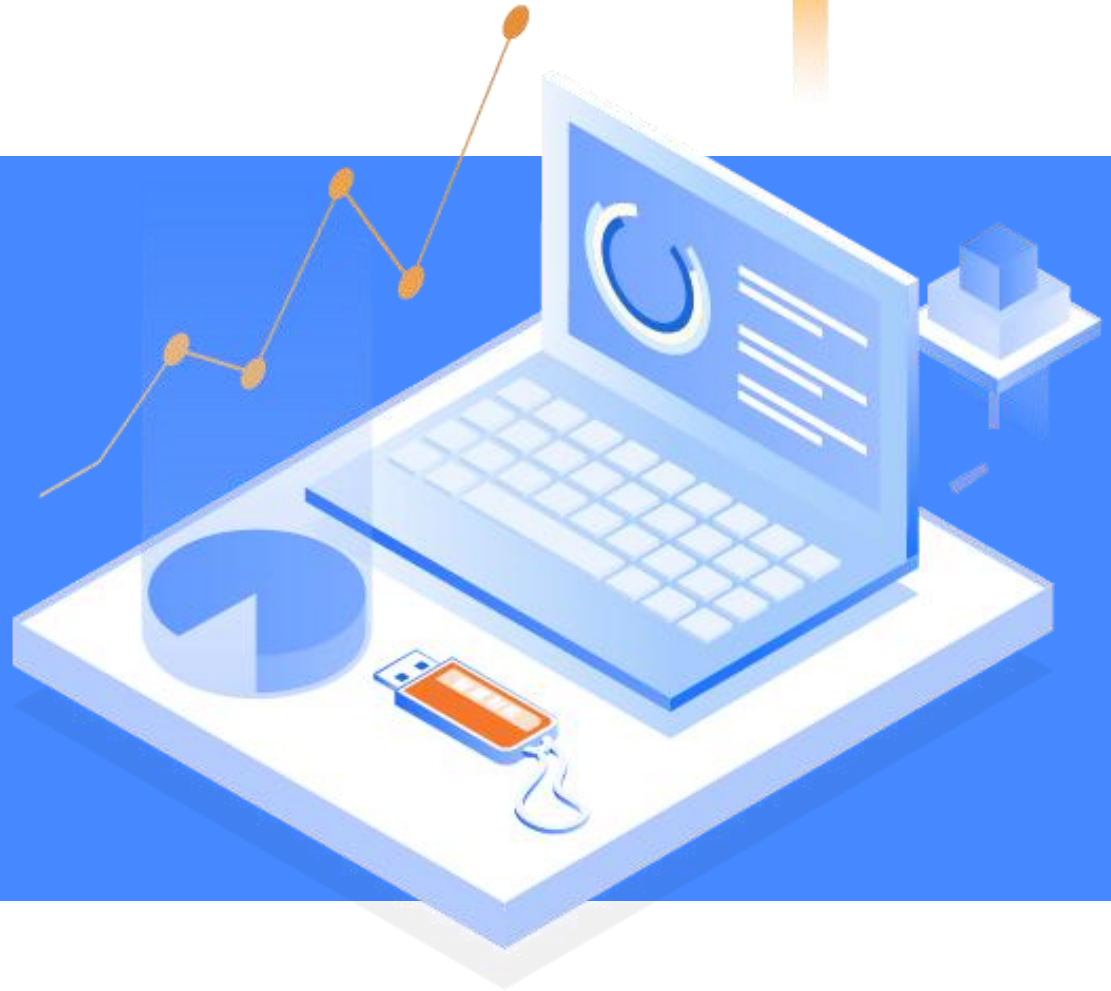


图7-6 选择模板配色

4.3 图片类AIGC应用实践

4.3.1 图片类AIGC应用场景
4.3.2 图片类AIGC案例实践



4.3 图片类AIGC应用实践

图片类AIGC是一种基于人工智能技术生成图片的方法，它利用深度学习、生成对抗网络（GAN）等先进算法，通过学习和模仿大量图像数据，能够自动创作出高度真实和艺术化的图片。AIGC在图像生成、修复、风格转换、艺术创作等领域展现出强大能力，为数字艺术、设计、游戏、电影等多个行业带来创新解决方案。其优势包括高效性、多样性和自动化，能够快速生成大量高质量的图像内容，满足各种复杂需求。

4.3.1 图片类AIGC应用场景

图片类AIGC的应用场景非常广泛，主要包括图像生成、图像修复、图像增强和图像识别等方面：

图像生成

AIGC能够生成高度逼真的图像，如人脸、动物、建筑物等。例如，OpenAI发布的DALL-E可以根据文本提示词创作出全新的、原创的图像，展示了AI在图像创作方面的强大能力。

图像修复

AIGC还可以修复损坏的图像，如去除噪声、填充缺失的部分等。这项技术对于保护和恢复古老的艺术作品、修复损坏的照片等具有重要意义。

图像增强

通过对图像进行增强处理，AIGC可以增加图像的饱满感和增强细节，使图像质量得到提升。这在提升照片的视觉效果、改善图像的清晰度和细节方面非常有用。

图像识别

AIGC在图像识别方面也有广泛应用，可以识别图像中的对象、场景和特征，如人脸识别、车牌识别等。这项技术对于安防监控、智能搜索、自动驾驶等领域的发展至关重要。

4.3.2 图片类AIGC案例实践

- 图片类AIGC大模型主要包括Midjourney、Stable Diffusion SDXL、百度文心一格等。这里以百度文心一格为例介绍图片类AIGC的使用方法。
- 文心一格是一款由百度公司研发的AI绘画工具，为用户提供了丰富的创意空间。使用文心一格进行AI绘画的步骤包括注册账户、选择创作模式、输入提示词、设置画面类型、设置比例、设置数量以及生成图片等，具体如下：

（1）注册账户：

访问文心一格官网（https://yige.baidu.com/），点击“注册”按钮，完成注册过程。

（2）选择创作模式：

进入文心一格首页以后，点击“立即创作”。在出现的界面中，在界面左上角位置选择“AI创作”，可供选择的模式包括推荐、自定义、商品图、艺术字、海报，可以满足不同的创作需求。这里可以选择默认的模式“推荐”。

4.3.2 图片类AIGC案例实践

（3）输入提示词：在提示词输入框中输入提示词，比如输入“请绘制一张图片，一个9岁的女孩子在海边沙滩上挖沙子”。

（4）设置画面类型：可以选择智能推荐、唯美二次元、中国风等各种类型。

（5）设置比例：可以选择竖图、方图、横图。

（6）设置数量：设置想要生成的图片的数量，比如设置为1。

（7）生成图片：点击“立即生成”，就可以生成相应的图片（如图所示）。图片生成以后，可以点击图片底部的“编辑本图片”，对图片进行编辑。

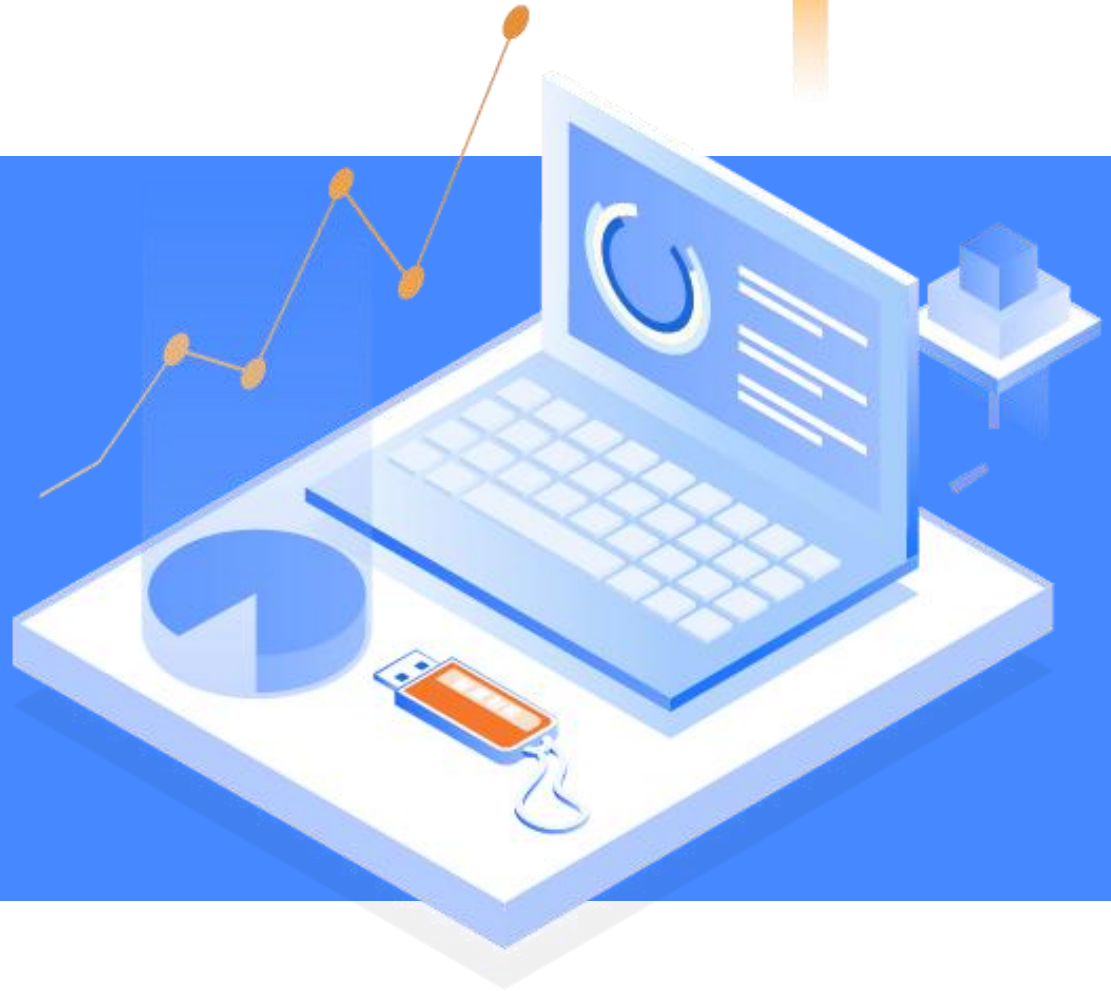
图 文心一格操作界面

4.3.2 图片类AIGC案例实践

文心一格提供了丰富的AI编辑功能，可以对图片进行各种智能化处理，包括图片扩展、图片变高清、涂抹消除、智能抠图、涂抹编辑、图片叠加等

4.4 语音类AIGC应用实践

4.4.1 语音类AIGC应用场景
4.4.2 语音类AIGC案例实践



4.4 语音类AIGC应用实践

语音类AIGC是一种利用人工智能技术（特别是语音识别、自然语言处理和语音合成技术），自动生成和处理语音内容的技术。它能够**模拟人类语音，实现语音到文本的转换、文本到语音的合成，以及语音情感分析等**功能，广泛应用于智能语音助手、智能客服、语音翻译等多个领域。

4.4.1 语音类AIGC应用场景

语音类AIGC的应用场景非常丰富，涵盖了多个领域，从日常生活到专业应用，都展现出了其独特的价值和潜力，以下是一些主要的语音类AIGC应用场景：

4.4.2 语音类AIGC案例实践

■ 豆包大模型的语音类功能用法

一般情况下，普通用户在手机上使用语音类AIGC大模型的场景比较多，因此，这里介绍手机版豆包的使用方法。

在智能手机上下载并安装“豆包APP”。启动进入豆包APP，会出现如图所示的对话界面，按住“语音按钮”（图中箭头指向的位置）不要松开，然后就可以对着手机说话，把自己的需求说出来，比如，可以说“请介绍一下厦门大学”，然后松开“语音按钮”，豆包就可以立即开始回答你提出的问题。豆包可以支持实时翻译，你可以语音输入“厦门大学的英文名称是什么”，豆包会马上给出翻译结果。

图 豆包的对话界面

4.4.2 语音类AIGC案例实践

■ **豆包大模型的语音类功能用法**

豆包不仅支持语音输入，也可以支持文字输入，只要在文字输入框内输入提示词，豆包就会给出回答。豆包也支持AI绘图功能，你可以用手指点击界面上的“图片生成”按钮，然后输入提示词，比如通过文字或者语音输入“请帮我绘制一张图片，一个9岁的小女孩在海边沙滩上玩沙子”，然后，豆包就会自动绘制生成满足你要求的图片

4.4.2 语音类AIGC案例实践

■ 豆包大模型的语音类功能用法

豆包还有一个很实用的功能，就是可以帮助你进行英语口语对话练习。

在豆包的操作界面的底部，用手指点击“对话”，在出现的功能选择界面中（如图所示），选择“英语口语聊天搭子”就可以进入英语口语聊天界面（如图所示），按住界面右下角的“语音按钮”，就可以开始用英语语音聊天了，你说完一句英语，松开语音按钮，豆包就会自动用英语语音回答你，然后你可以继续输入语音进行后续对话。

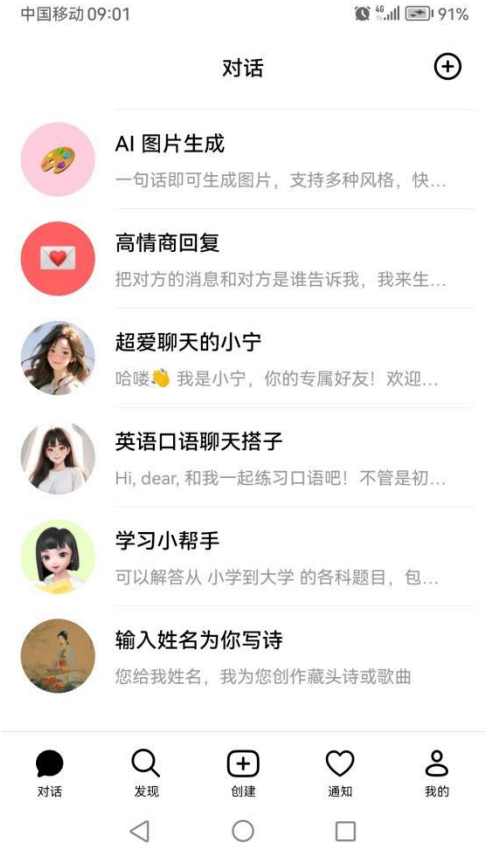


图 豆包的功能选择界面

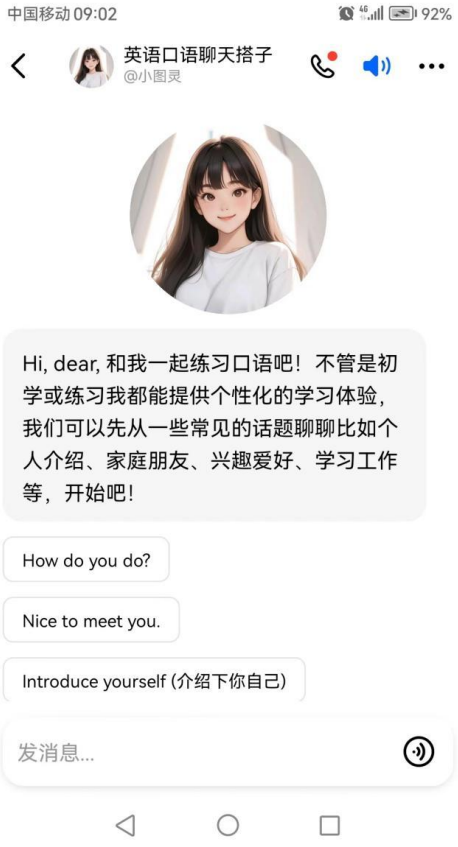


图 豆包的英语口语聊天界面

4.4.2 语音类AIGC案例实践

■ 讯飞智作大模型的语音类功能用法

访问讯飞配音官网（https://peiyin.xunfei.cn/），首先按照页面提示完成用户注册。注册成功以后，会进入“讯飞智作”页面（如图4-13所示），在页面顶部选择“讯飞配音”。

在讯飞配音页面（如图4-14所示），输入你的配音文本内容，比如输入“人工智能是新一轮科技革命和产业变革的重要驱动力量，是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学”。可以设置配音的品质，点击页面左上角“叙述（品质）”，在出现的页面中（如图4-15所示），可以选择自己喜欢的主播类型，并且允许对主播的语速和语调进行设置，然后再点击页面右上角的“使用”按钮。然后，点击页面右上角的“生成音频”，在出现的页面中（如图4-16所示），设置作品名称、文件格式和字幕，再点击“确认”。

这时，会出现订单支付页面（如图4-17所示），可以选择“会员及语音包购买”（45元一个月）或者“单次付费”（20元每次）。完成费用支付以后，就会出现下载提示页面（如图4-18所示），点击“去下载”，然后，在出现的下载页面中（如图4-19所示），点击下载按钮（图中箭头指向的位置）就可以把配音文件下载到本地电脑中。在本地电脑播放下载后的配音文件可以发现，现在的AI配音技术已经比较成熟，生成的配音质量已经可以达到专业配音员的水平。

4.4.2 语音类AIGC案例实践

■ 讯飞智作大模型的语音类功能用法

图4-13 讯飞智作页面

图4-14 讯飞配音页面

4.4.2 语音类AIGC案例实践

■ 讯飞智作大模型的语音类功能用法

图4-15 选择主播

4.4.2 语音类AIGC案例实践

■ 讯飞智作大模型的语音类功能用法

图4-16 作品命名页面

图4-17 订单支付页面

4.4.2 语音类AIGC案例实践

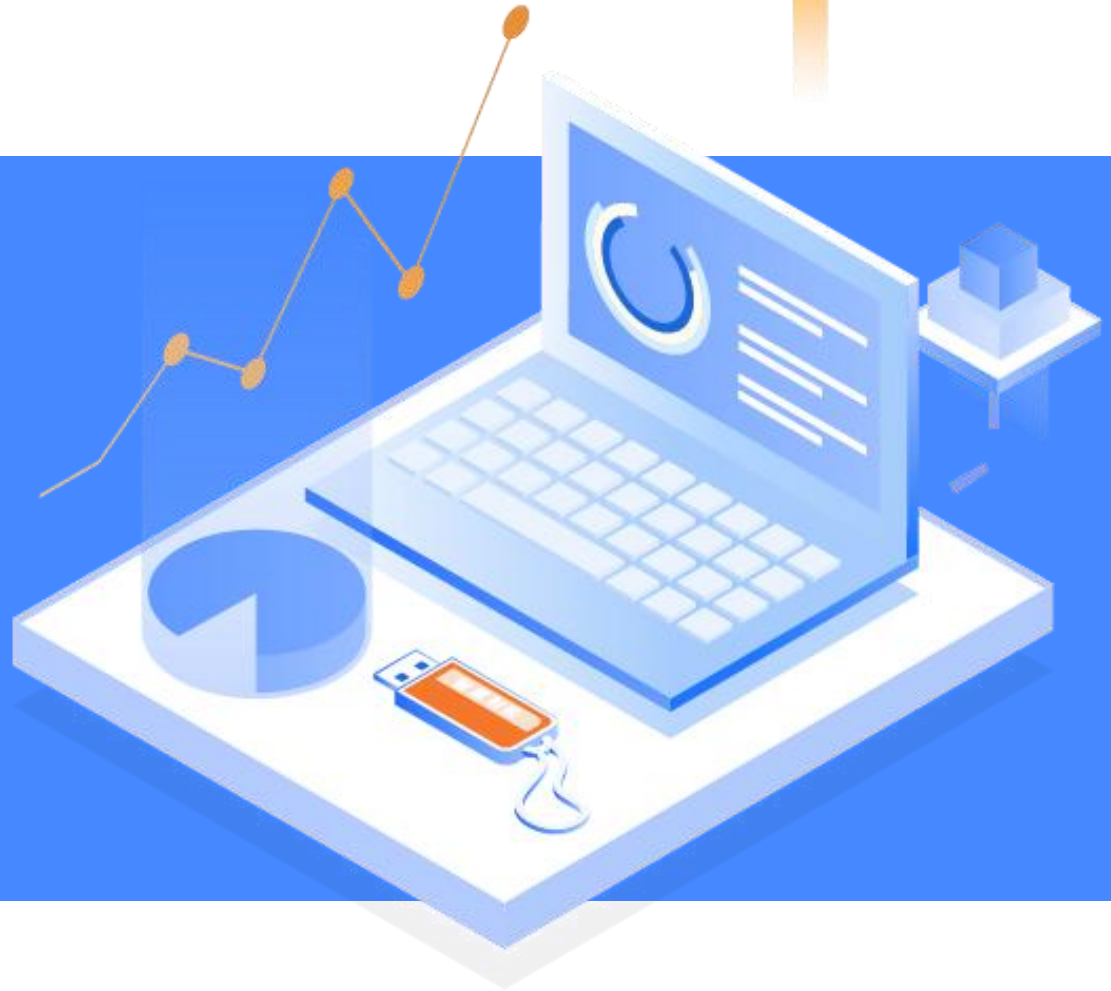
■ 讯飞智作大模型的语音类功能用法

图4-18 下载页面

图4-19 点击下载按钮

4.5 视频类AIGC应用实践

4.5.1 视频类AIGC应用场景
4.5.2 代表性视频类AIGC大模型
4.5.3 视频类AIGC案例实践



4.5视频类AIGC应用实践

视频类AIGC是指利用人工智能技术，特别是深度学习、机器学习等算法，自动创建或处理视频内容的技术。**它能根据给定的文本、图像或其他数据，自动生成符合描述的视频内容，涵盖文生视频、图生视频、视频风格化、人物动态化等多个方向**。这一技术在创意设计、影视制作等领域潜力巨大，极大地提升了视频内容的生产效率和质量

4.5.1 视频类AIGC应用场景

视频类AIGC在多个领域拥有广泛的应用场景，以下是一些主要的应用方向：

- 影视制作与后期制作
- 短视频与直播
- 广告与营销
- 教育与培训
- 虚拟现实与增强现实
- 新闻传播与媒体融合
- 其他领域

4.5.2 代表性视频类AIGC大模型

- 视频类AIGC大模型发端于Sora。2024年2月，美国的Open AI发布了全球第一款文生视频大模型Sora（这里的“文生视频”是指由输入的文本内容生成相应的视频），迅速引起了业界的广泛关注和讨论，因其能够快速生成高质量的广告宣传视频及商品演示视频，从而大幅降低广告相关内容的制作成本及时间。
- 我国的视频类AIGC大模型主要包括：

可灵

由快手推出，被誉为中国版Sora，视频生成时长可达120秒，支持文生视频、图生视频、视频续写、镜头控制等功能，表现出色。

Vidu

生数科技联合清华大学发布，是中国首个长时长、高一致性、高动态性视频大模型，支持一键生成16秒高清视频，性能对标国际顶尖水平。

书生·筑梦

由上海人工智能实验室研发，可生成分钟级视频，已用于央视AI动画片《千秋诗颂》的制作，具备中国元素和高清画质。

由于视频类AIGC大模型在使用时，会消耗大量的算力资源，使用成本很高，所以，目前国内的视频类AIGC大模型大多数没有免费开放给大众使用，即使是免费使用，也只能生成很短时间长度的视频。

4.5.3 视频类AIGC案例实践

案例：使用腾讯智影生成数字人播报视频（备注：本案例由夏小云老师制作）

步骤1：登录腾讯智影平台。在浏览器地址栏中输入网址“https://zenvideo.qq.com/”，进入“腾讯智影”平台，点击“登录”（如图7-107所示），可以使用微信扫码登录、也可以手机号登录或者QQ扫码登录，任选一种方式登录即可。登录成功后，点击平台首页“智能小工具”栏目中的“数字人播报”按钮（如图7-108所示）或者“智能小工具”上方的“数字人播报”按钮，进入“数字人播报”功能界面。

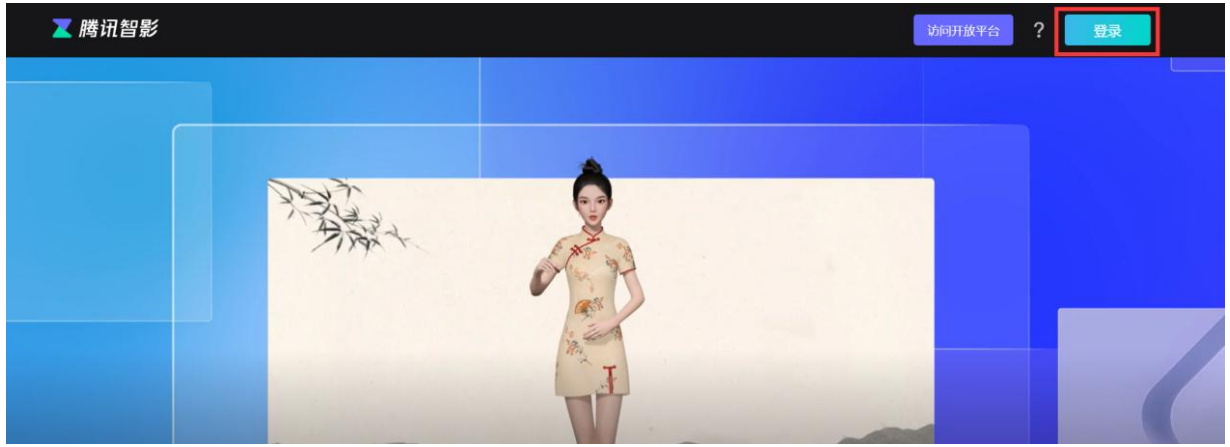


图7-107 腾讯智影登录界面

图7-108 “数字人播报”功能入口

4.5.3 视频类AIGC案例实践

步骤2：上传PPT。进入“数字人播报”功能界面后（如图7-109所示），在左侧工具栏，点击“PPT模式”,平台会出现“上传PPT或PDF”的界面，点击“上传”按钮，上传需要播报的PPT，这里上传“数字人播报PPT.ppt”文件（可以从林子雨编著《数字素养通识教程》教材官网下载）。

图7-109 “PPT模式”上传PPT入口

4.5.3 视频类AIGC案例实践

步骤3：选择数字人。PPT上传结束后，点击左侧工具栏“数字人”按钮，会出现“数字人”设置界面，包括“预置形象”和“照片播报”两大板块（如图7-110所示）。“预置形象”分为“2D数字人”和“3D数字人”；“照片播报”分为“照片主播”和“AI绘制主播”两种（如图7-111所示）。这里选择“预置形象”里面的“2D数字人”——卓妤数字人，作为PPT播报的数字人。

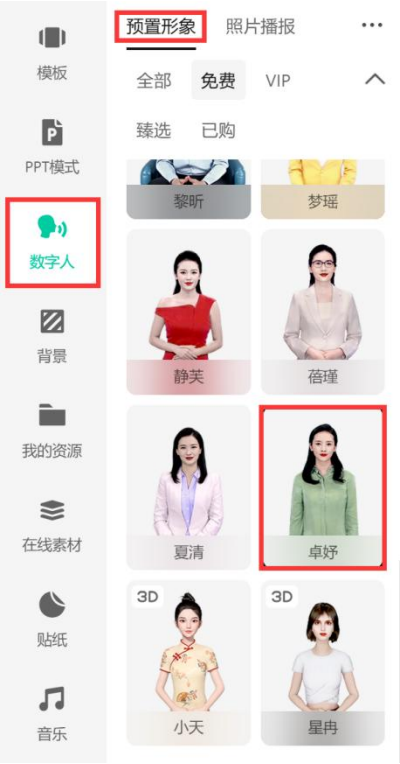


图7-110 “预置形象”的数字人

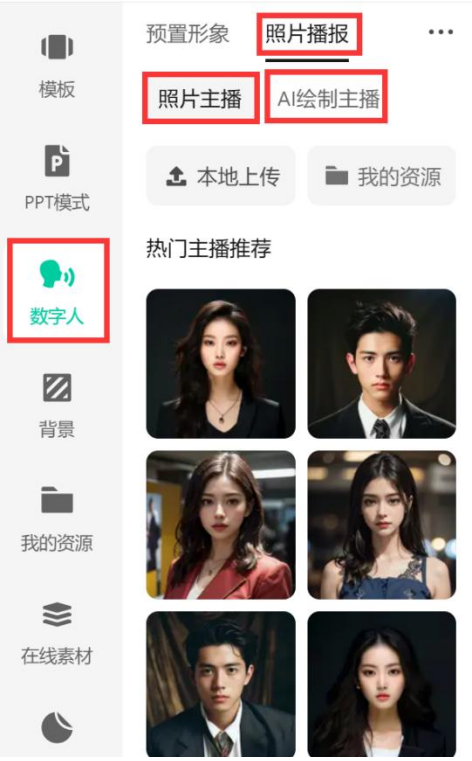


图7-111 “照片播报”的数字人

4.5.3 视频类AIGC案例实践

步骤4：调整数字人的位置、大小和服装类型。选用的数字人确认后，可以点击PPT上的“数字人”，进入“数字人”设置界面，点击“数字人编辑”按钮（如图7-112所示），对“数字人”进行编辑，可以重新换服装以及选择数字人出现的“形状”；点击“画面”按钮，可以通过坐标设置来调整数字人的位置和大小（如图7-113所示），也可以点击数字人的边框，通过拖动鼠标来调整数字人的位置和大小。这里可以根据PPT画面的布局来调整“数字人”的位置和大小，尽量避免数字人遮挡文字。

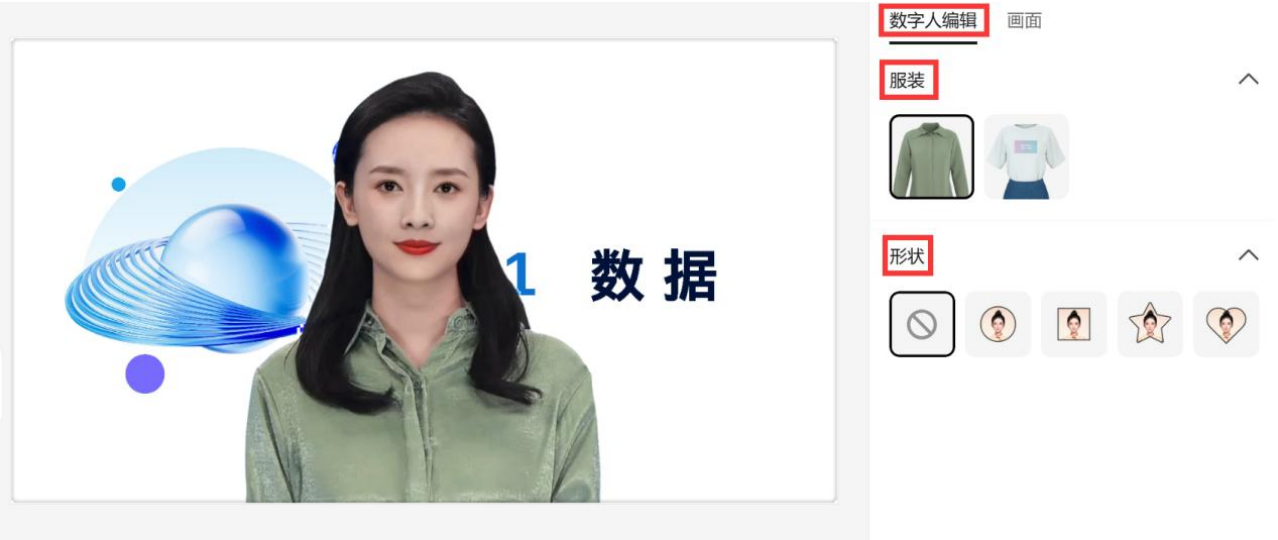


图7-112 “数字人编辑”功能界面确认服装和形状

图7-113 进入“画面”功能界面调整数字人的位置和大小

4.5.3 视频类AIGC案例实践

步骤5：输入播报内容和设置字幕样式。数字人调整完成后，点击右侧工具栏的“播报内容”按钮（如图7-114所示），输入播报内容，播报内容可以AI自动生成，也可以手动输入或导入文件（字数不超过5000字），这里选择“手动输入”每页PPT的播报内容。点击右侧工具栏的“字幕样式”按钮（如图7-115所示），设置视频字幕的样式并打开字幕显示按钮，通过鼠标拖动把字幕放置画面的合适位置。

图7-114 输入PPT播报内容

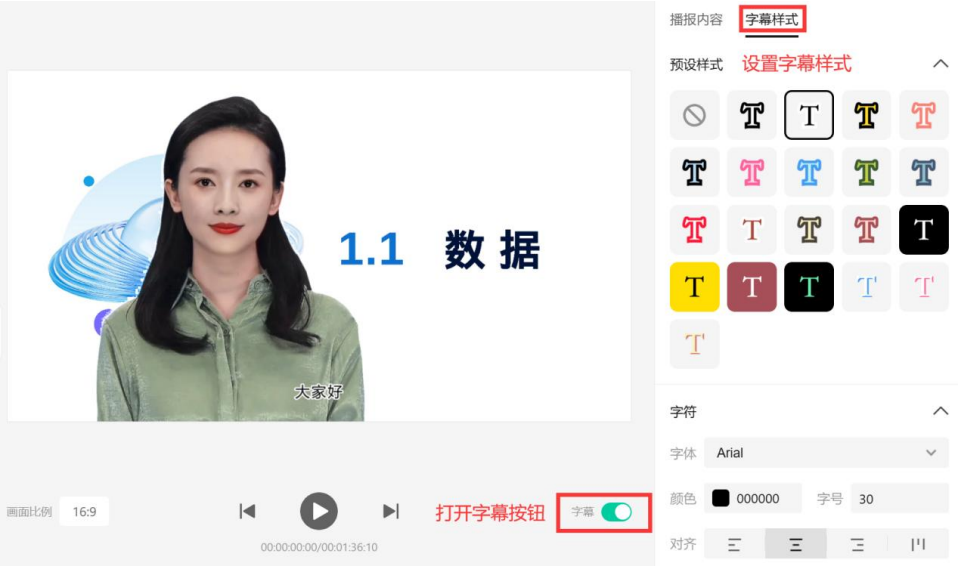


图7-115 设置字幕样式并打开字幕按钮

4.5.3 视频类AIGC案例实践

步骤5：根据需要添加“背景”、“贴纸”、“音乐”并选择播报音色。输入完播报内容后，可以点击左侧工具栏的“背景”、“贴纸”和“音乐”等功能（如图7-116所示），根据需要添加，这里选择不添加背景、贴纸和音乐。确认后，点击“播报内容”输入框下方的“音色”按钮，选择合适的音色。这里选择“如云”音色作为PPT的数字人播报音色（如图7-117所示）。

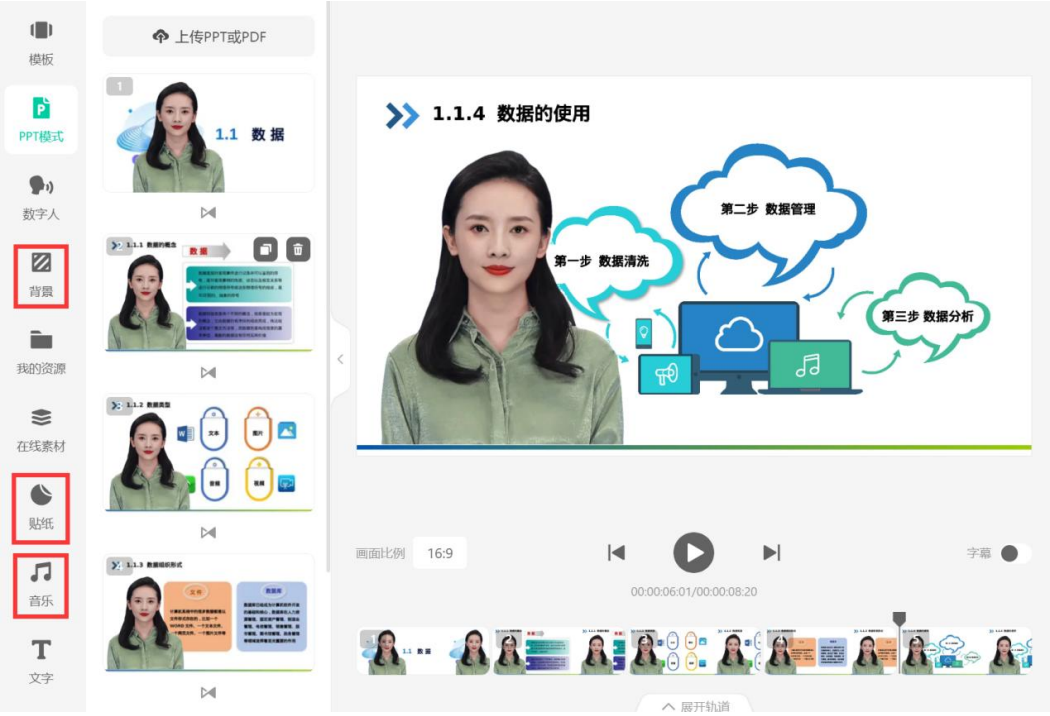


图7-116 根据需要添加“背景”、“贴纸”、“音乐”

图7-117 选择播报音色

4.5.3 视频类AIGC案例实践

步骤6：保存并生成播报。选择完音色后，点击“音色”下方的“保存并生成播报”按钮（如图7-118所示），选中每一页PPT，逐一点击保存并生成播报，注意，平台此时生成的数字人播报效果预览，暂不支持口型对齐预览，合成后可查看完整动态效果。

图7-118 保存并生成播报

4.5.3 视频类AIGC案例实践

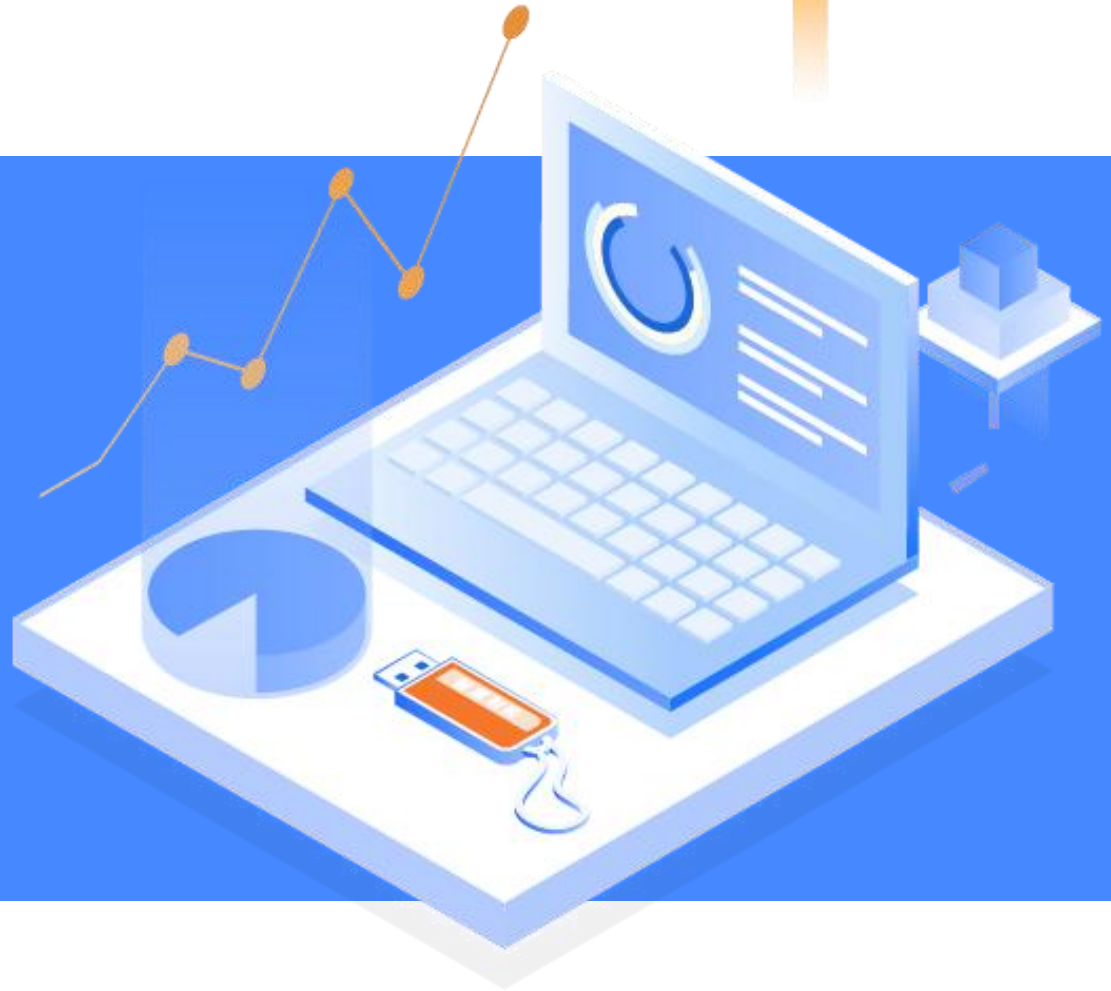
步骤7：合成并下载视频。保存并生成每页的PPT的播报后，点击页面右上方的“合成视频”按钮（如图7-119所示），然后设置合成视频输出的参数（如图7-120所示），设置完成后点击“确认”按钮，系统后台会自动合成数字人播报视频，等待合成结束后，点击“下载”按钮（如图7-121所示），下载合成的数字人播报视频。最后，播放合成的视频文件，检查视频画面是否符合预期。如有需要，可以根据反馈调整参数，重新生成。

图7-119 点击合成视频

图7-120 合成视频参数设置

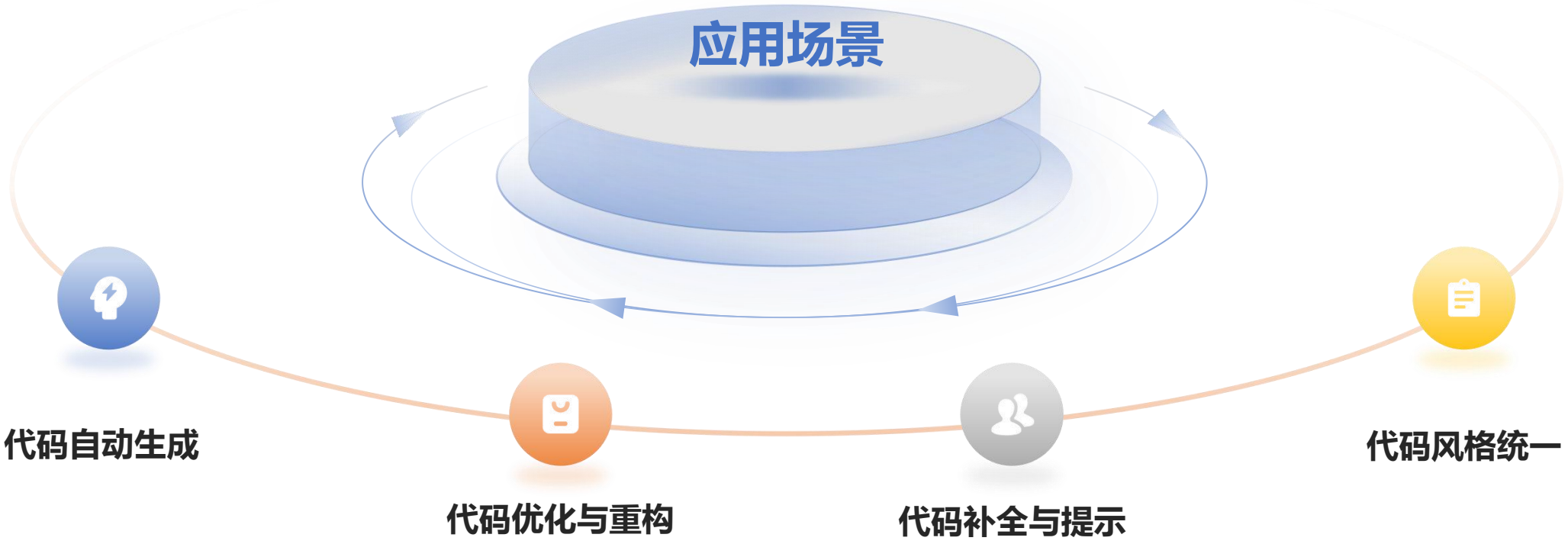
图7-121 下载数字人播报视频

4.6 AIGC在辅助编程中的应用



4.6 AIGC技术在辅助编程中的应用

AIGC技术在辅助编程中的应用日益广泛，它能够自动生成高质量的代码，从而显著提高开发效率，主要包括以下几种应用场景：



能够提供辅助编程服务的AIGC大模型包括Codex、GitHub Copilot、CodeGeeX、aiXcoder、豆包、通义灵码等。这里以我国的字节跳动公司研发的豆包大模型为例介绍使用方法。

4.6 AIGC技术在辅助编程中的应用

编程工作一般是在电脑上进行，所以这里使用电脑端的豆包大模型（手机端的豆包APP也提供了编程辅助功能）

访问豆包大模型官网（https://www.doubao.com/），注册用户以后，进入大模型操作首页（如图所示），点击“我的智能体”，再点击“编程助理”，然后，在页面中输入提示词，比如输入“请编写一段Python代码，使用turtle库，绘制一个五角星”，然后，豆包就会自动生成一段Python代码（如图所示）。在Python中运行这段代码，就可以成功绘制一个五角星。

图 豆包大模型操作首页

4.6 AIGC技术在辅助编程中的应用

请编写一段Python代码，使用turtle库，绘制一个五角星

以下是一段使用 Python 的 `turtle` 库绘制五角星的代码：

收起

python 复制

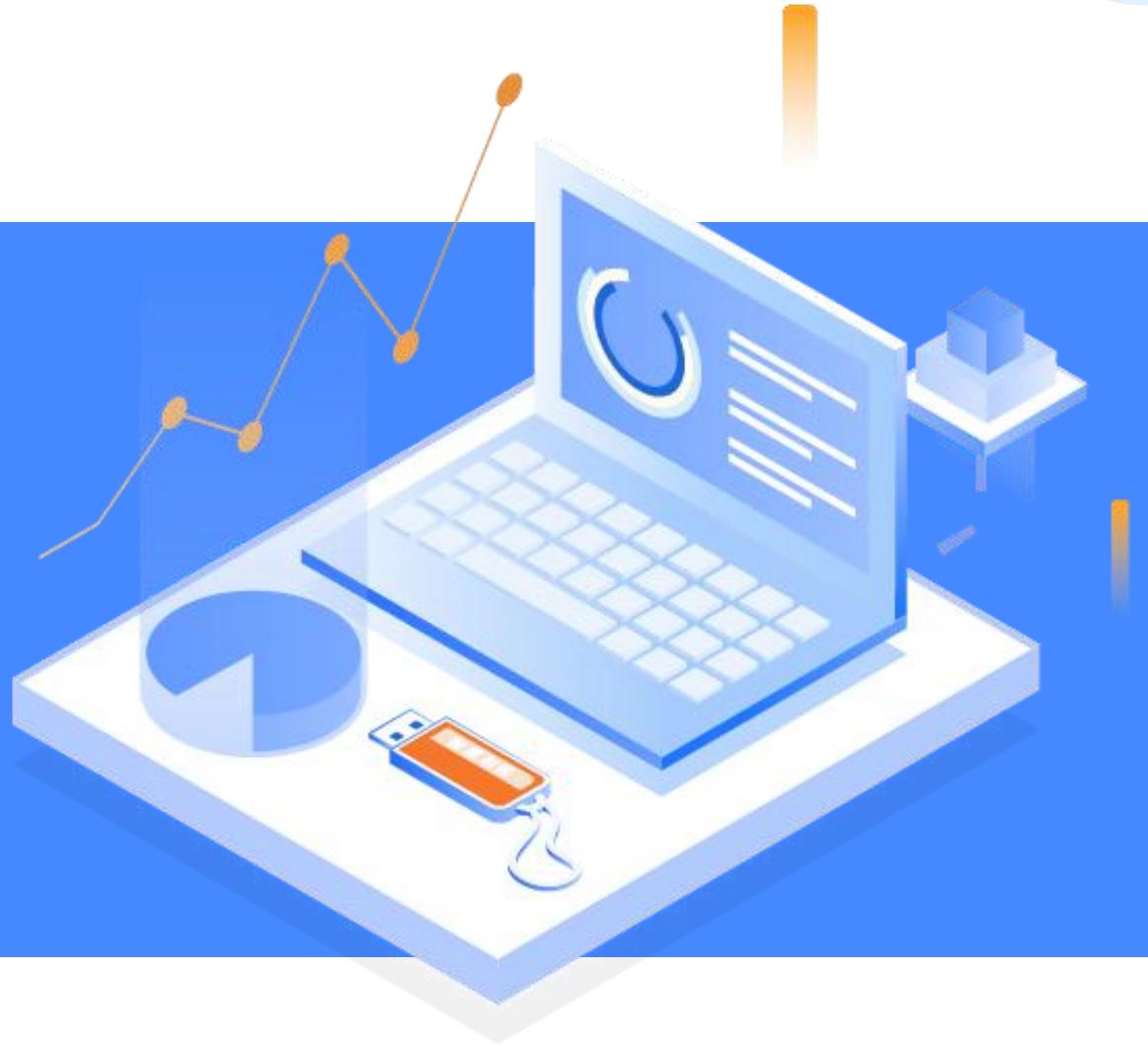
```python
import turtle

# 绘制五角星
def draw_pentagon():
    turtle.right(72)
    for _ in range(5):
        turtle.forward(200)
        turtle.right(144)

draw_pentagon()
turtle.done()
```

这段代码首先导入了 `turtle` 库。然后定义了一个名为 `draw_pentagon` 的函数，在函数内部，通过不断调整角度和前进距离来绘制五角星的五条边。最后使用 `turtle.done()` 来保持图形窗口的显示。

复制 重新生成 ···

图 豆包自动生成的Python代码

4.7 AI搜索



4.7 AI搜索

AI搜索，即人工智能搜索引擎，是一种利用先进的人工智能技术，特别是深度学习和自然语言处理（NLP），来理解和响应用户的查询需求的新型搜索工具。它不仅仅是传统搜索引擎（比如百度）的简单升级，而是通过模拟人类的思维方式和行为模式，为用户提供更加精准、个性化且高效的信息检索服务。AI搜索通过收集和分析用户的历史搜索数据和行为模式，构建用户画像，从而实现更加精准的个性化搜索服务。这种数据驱动的智能决策机制，使得AI搜索能够不断自我优化，提升用户体验。

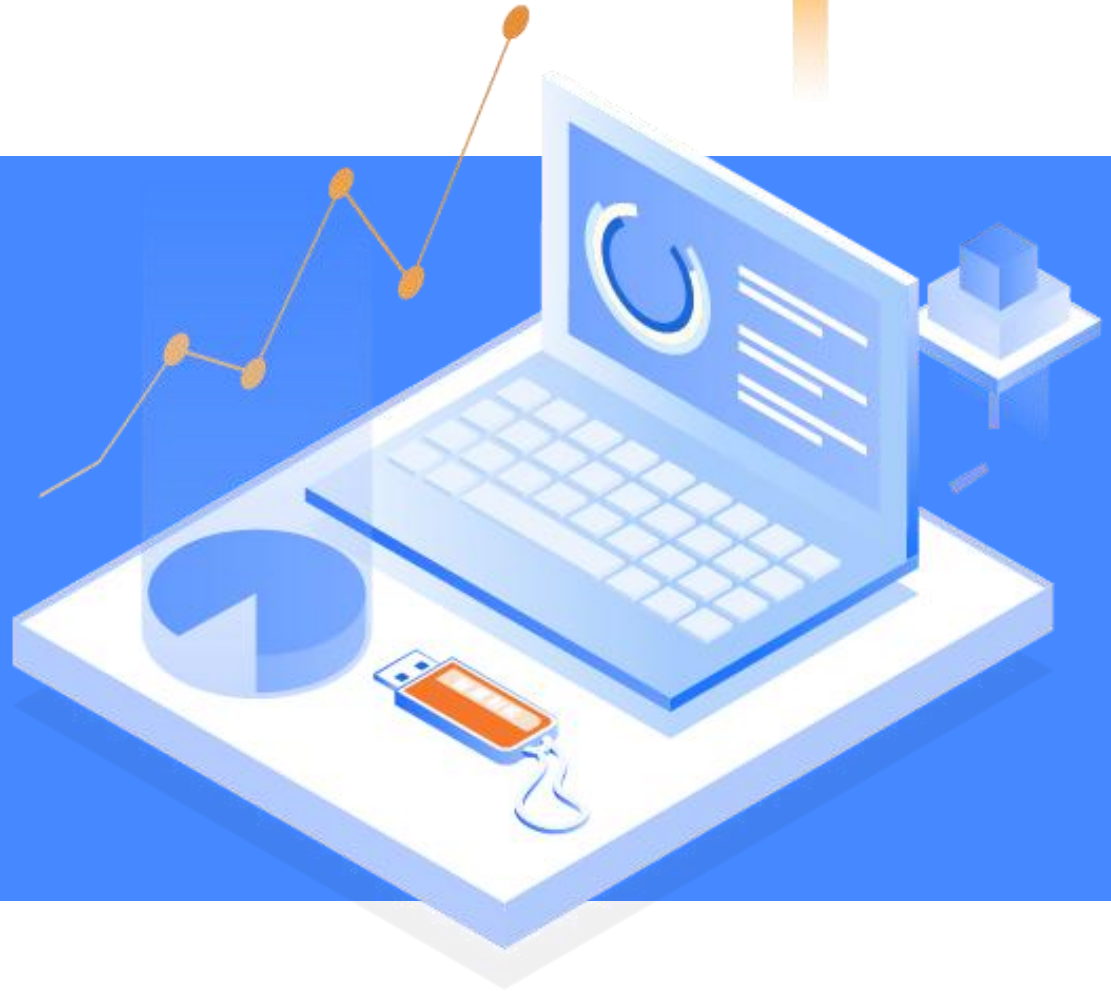
纳米AI搜索是360公司在2024年12月推出的全新AI搜索应用，结合了自然语言处理、机器学习以及专家协同技术，致力于打破传统搜索引擎的局限，提供智能化、多样化的搜索体验。其核心特点包括：

（1）多模态搜索：支持文字、语音、拍照、视频等多种输入方式，满足不同场景下的需求，实现“一切皆可搜索”。

（2）智能工具集成：内置16款顶尖大模型，如豆包、文心一言等，为用户提供一站式AI智慧体验。

（3）慢思考模式：通过专家协同和多模型协作，深入分析复杂问题，提供更专业、更全面的答案。

4.8 AI智能办公



4.8 AI智能办公

在人工智能时代，AI智能办公正以前所未有的态势重塑我们的工作模式与体验，成为推动办公效率提升和办公方式变革的核心力量。
AI智能办公将人工智能技术深度融入办公场景的各个环节，例如文档处理、数据分析、演示制作等。

在**文档处理**领域，AI智能办公带来了前所未有的变革。以往需要人工手动输入文字、排版以及校对纠错等工作需要耗费大量时间和精力，如今AI文档处理技术极大地提升了这些任务的效率与质量。具体如下：
（1）生成式文本。通过对大量文本数据的学习，AI能够根据用户提供的提示词、主题或简单描述，快速生成内容完整、逻辑连贯的文档初稿。无论是新闻稿件、公告通知还是学术论文，都能借助这一功能节省撰写时间。例如，市场调研公司在需要撰写季度报告时，AI可以迅速整合数据和市场趋势信息，生成报告框架与初步内容，供使用者进一步完善。
（2）文档智能排版。AI能够自动识别文档内容的结构，如标题、段落、列表等，并根据用户预设的格式模板进行快速排版。这不仅提高了排版效率，还确保了文档格式的一致性和规范性。对于需要处理大量文档的办公人员来说，这一功能大大减轻了工作负担。

4.8 AI智能办公

数据分析是办公场景中的重要环节。传统的数据分析需要人工进行数据收集、清洗、分析和可视化，过程烦琐且容易出错。而借助人工智能技术，这些工作可以更加高效、精准地完成。

通过大模型算法，AI能够从海量数据中发现潜在的模式、趋势和关联关系。例如，电商企业可以利用AI分析用户的购买行为、浏览记录和搜索关键词，挖掘出用户的潜在需求和消费偏好，从而制定精准的营销策略。

自动数据可视化功能让数据分析结果的呈现更加直观、清晰。AI能够根据数据分析结果自动生成各种类型的表格图表和图形，如柱状图、折线图、饼图等，并进行合理的布局和配色。办公人员无需花费大量时间手动制作图表，就能快速将数据转化为易于理解的可视化信息，为决策提供有力支持。

4.8 AI智能办公

WPS Office

AI写作助手（帮我写、帮我改、AI伴写）
AI设计助手（AI排版、AI格式）
AI阅读助手（全文总结、文档问答、划词解释和翻译）
AI数据助手（AI写公式、AI数据分析）

4.9 总结

大模型是人工智能领域的重要研究方向，其强大的语言理解和生成能力使得它在自然语言处理、机器翻译、智能客服等领域有着广泛的应用。大模型的训练需要大量的数据和计算资源，同时也需要先进的技术和算法支持。随着技术的不断发展，大模型的应用场景也在不断扩展，未来将会更加广泛地应用于各个领域。

谢谢