

AI 编程与智能体开发

林子雨 副教授 厦门大学
E-mail: ziyulin@xmu.edu.cn
2026年9月



主讲教师



- ◆ 2013年度厦门大学奖教金获得者
- ◆ 2017年度厦门大学奖教金获得者
- ◆ 2020年度厦门大学奖教金获得者
- ◆ 2023年度厦门大学奖教金获得者
- ◆ 2026年度厦门大学奖教金获得者
- ◆ 2018年国家精品在线开放课程
- ◆ 2020年国家级线上一流本科课程
- ◆ 2023年国家级线上一流本科课程
- ◆ 2020年福建省线上一流本科课程
- ◆ 2021年福建省线上线下混合一类本科课程
- ◆ 2018年福建省高等教育教学成果奖二等奖
- ◆ 2022年福建省高等教育教学成果奖特等奖
- ◆ 2024年福建省高等教育教学成果奖特等奖
- ◆ 2023年教育部国家智慧教育公共服务平台应用典型案例





配套教材



《AI编程与智能体开发》

林子雨 编著

人民邮电出版社 预计2026年11月上市销售

教材官网提供详细信息和样书申请

官网: <https://dbllab.xmu.edu.cn/post/ai-programming-agent/>

教材目录

- 第1章 大模型
- 第2章 AI编程概述
- 第3章 AI编程环境搭建
- 第4章 基于提示词的AI编程
- 第5章 基于规范的AI编程
- 第6章 流程驱动的AI编程
- 第7章 智能体概述
- 第8章 智能体开发框架LangChain
- 第9章 智能体开发框架LangGraph

教材特色

- 教材内容来自互联网大厂真实实践
- 以实践教学为依托
- 理论实践贯通
- 覆盖主流技术



扫码访问教材官网
了解详情、获取资源、申请样书

第1章 大模型



目录

01 大模型概述

02 大模型产品

03 大模型的基本原理

04 大模型的特点

05 大模型的分类

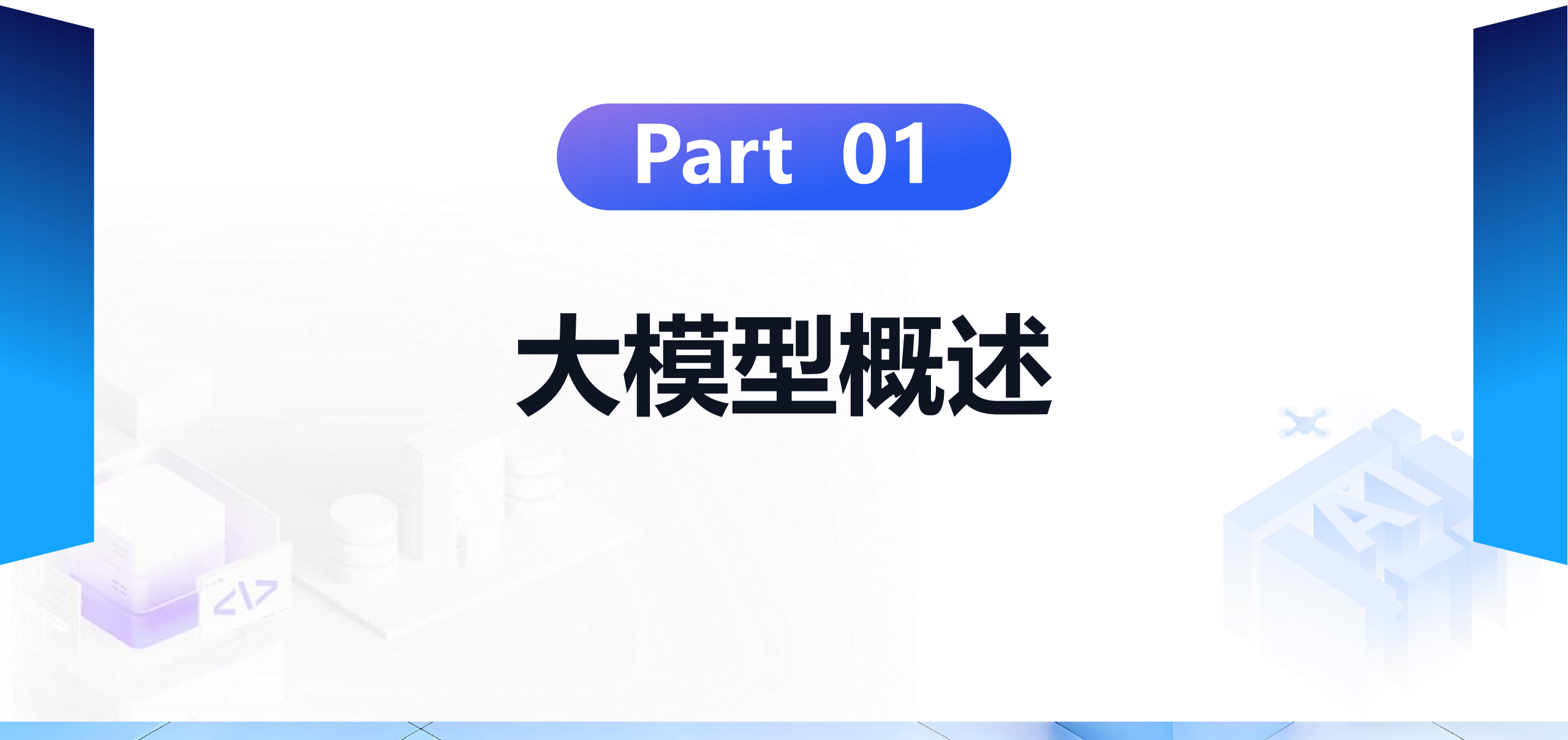
06 大模型的应用

07 大模型相关技术与概念的演化



Part 01

大模型概述





1.1 大模型概述



大模型的概念

大模型与
小模型的区别

大模型
的发展历程

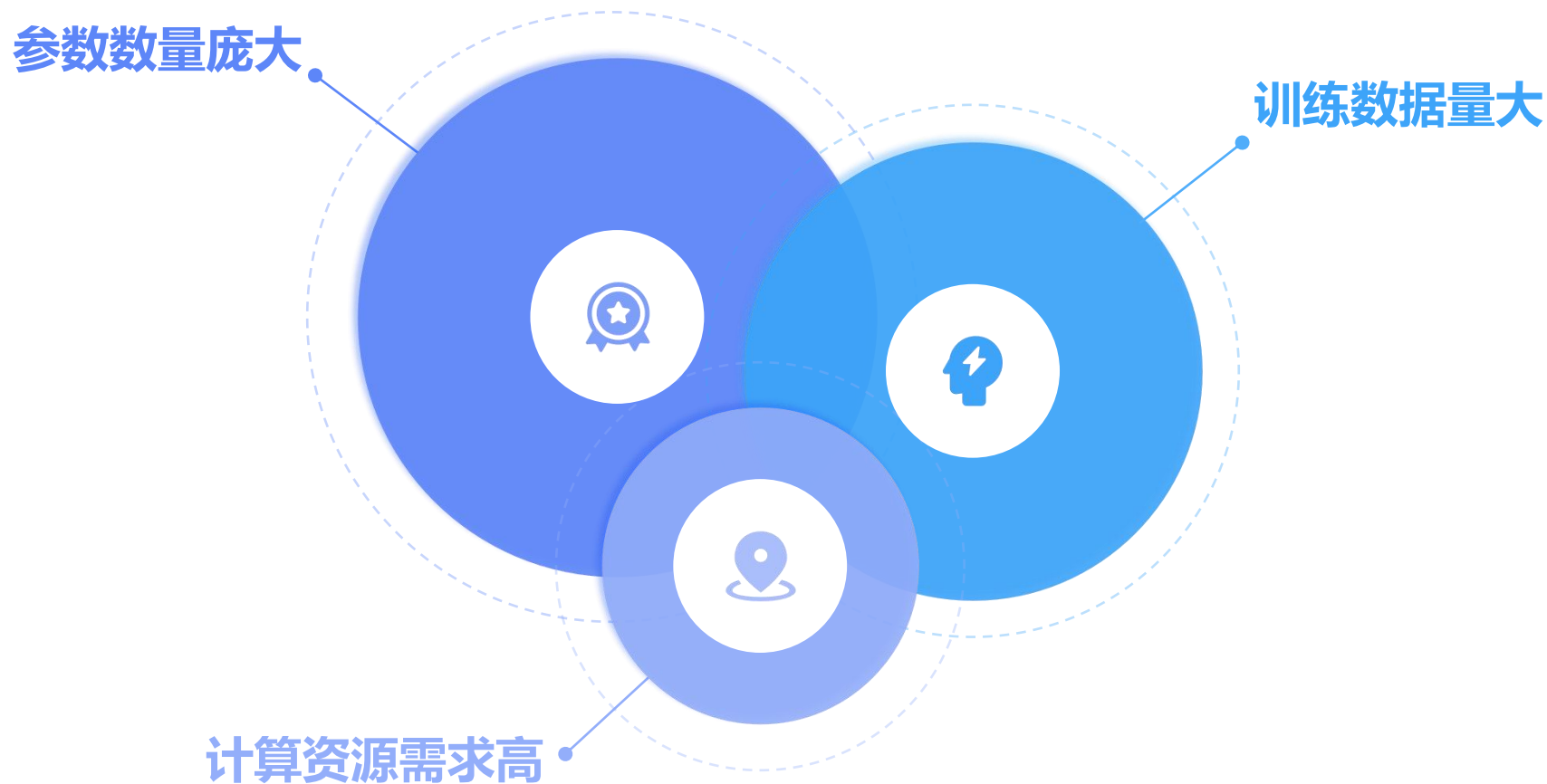
人工智能与
大模型的关系



1.1.1 大模型的概念



大模型通常指的是**大规模的人工智能模型**，它基于深度学习技术，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





1.1.2 大模型与小模型的区别

01 小模型定义、特征与适用场景

参数少、网络层数浅；优势为轻量化、推理高效、部署简单；适配移动端、嵌入式、物联网等算力、数据有限场景

02 大模型核心判定标准：涌现能力

模型参数、训练数据达到临界规模后，自动自主学习高阶特征与全新复杂能力；拥有涌现能力是大模型与小模型最本质区别

03 大模型基础特点与适用场景

参数多、网络层次深，表达力、任务精度更强；缺陷是训练、推理消耗大量算力与时间；适配云端、高性能计算等充足算力、大数据场景

04 小模型阶段 AI4S 的局限

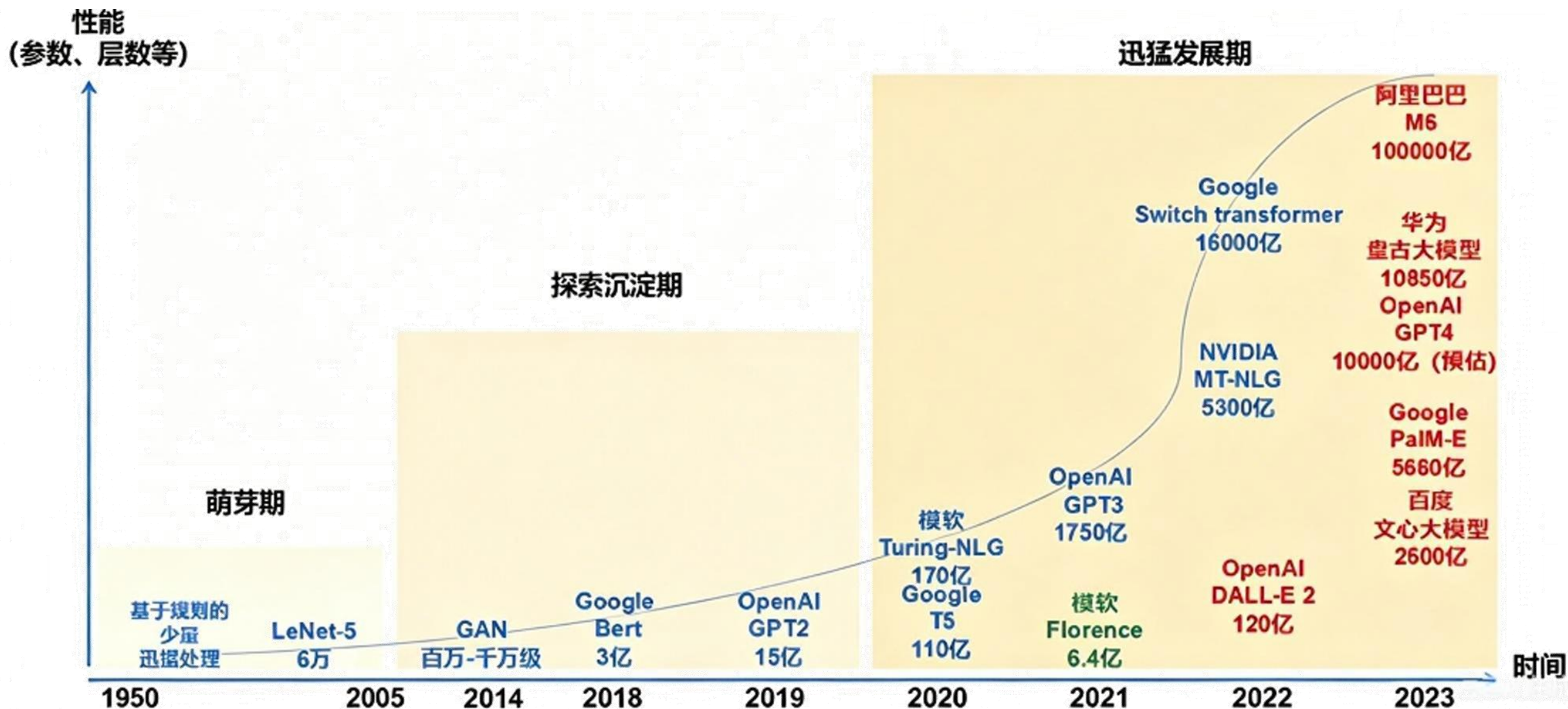
传统小模型 AI 仅针对单一科研任务优化，只能依托已知规律做数据处理、简单预测；泛化能力弱，无法处理高复杂度综合性科学问题

05 大模型时代 AI4S 的核心赋能优势

①海量数据支撑跨领域知识融合；②可捕捉高维复杂数据结构，实现科学问题全局综合分析；③生成式能力可产出创新科研假设，拓展科研新思路

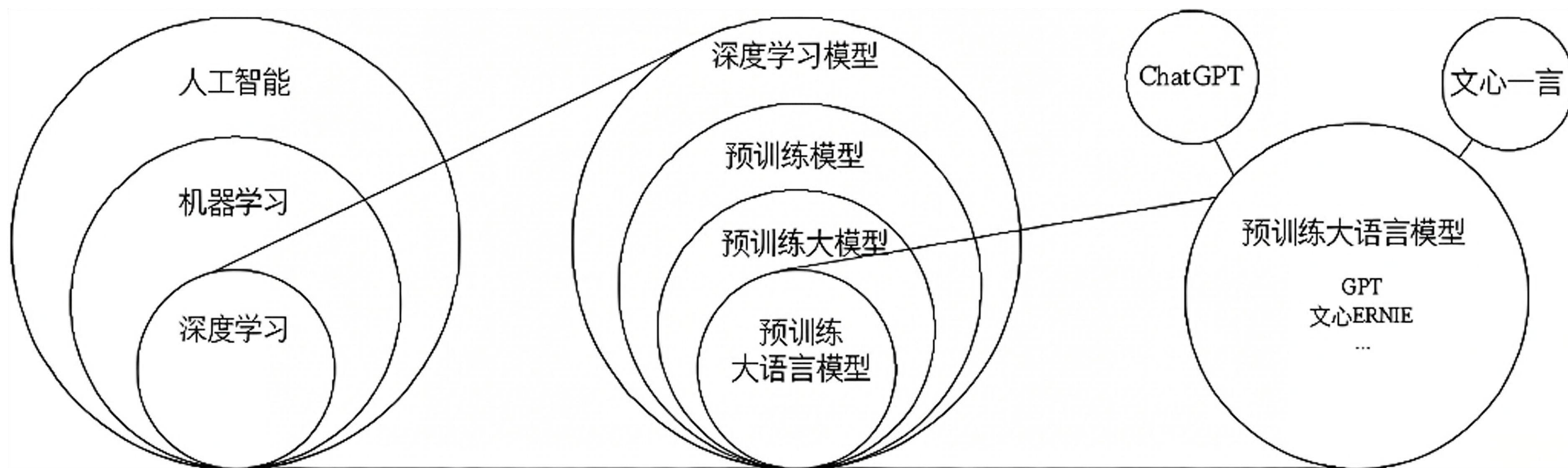


1.1.3 大模型的发展历程



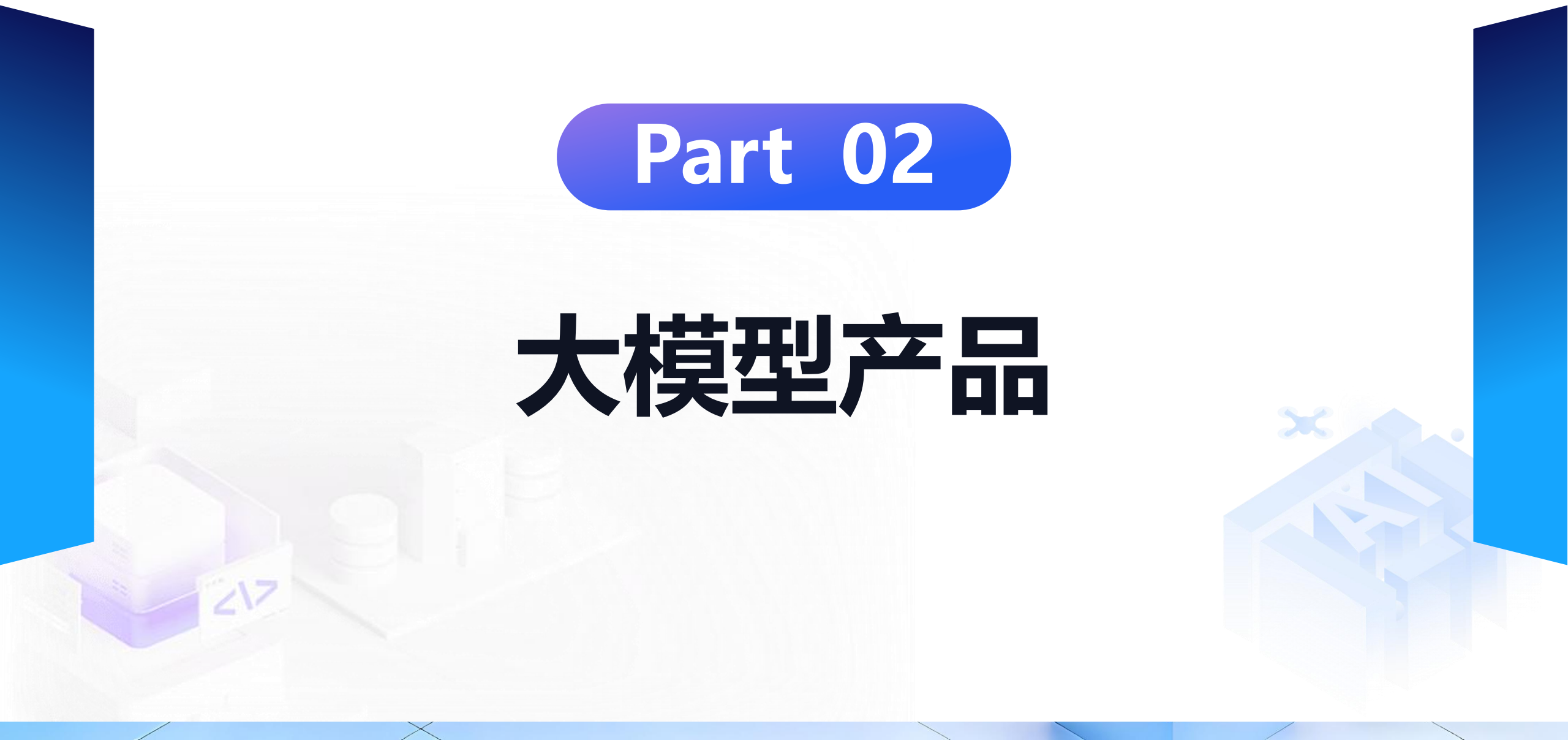


1.1.4 人工智能与大模型的关系



Part 02

大模型产品





1.2 大模型产品



国外的
大模型产品

国内的
大模型产品



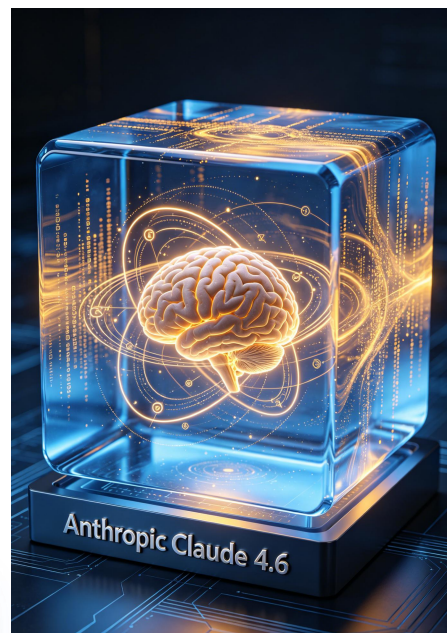
1.2.1 国外的大模型产品



OpenAI GPT-5.6



Google Gemini 3



Anthropic Claude
4.8



Meta Llama 4



xAI Grok 4.3



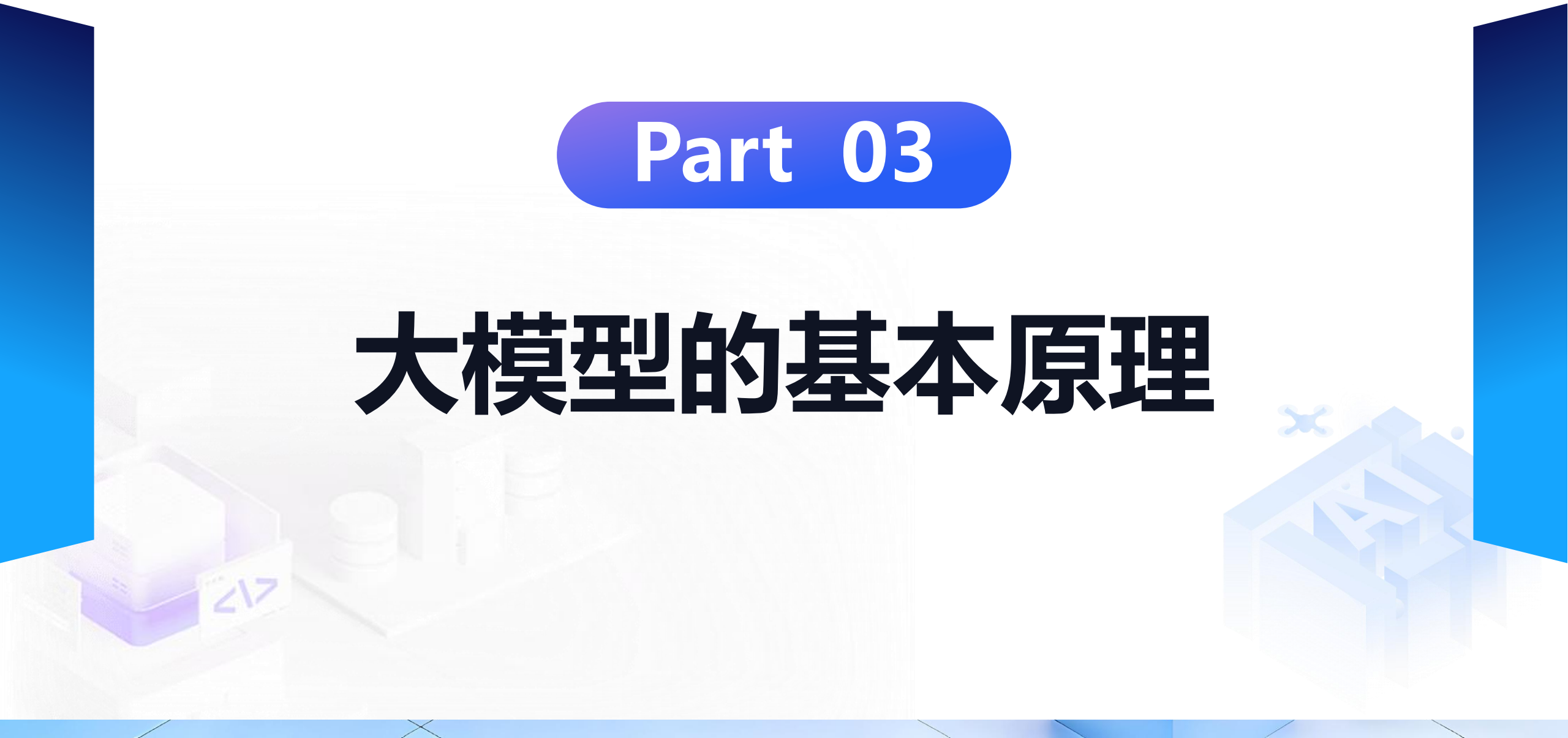
1.2.2 国内的大模型产品



大模型	图标	指标排名
DeepSeek		能力测评第一，用户数第二
豆包		用户数量第一，日活用户数第一
Kimi		文本处理第一
即梦AI		作图能力第一(Seedream5.0) 视频生成第一(Seedance2.0)
智谱清言		编程能力第一

Part 03

大模型的基本原理





1.3 大模型的基本原理



原理概述

大模型训练的
实例演示



1.3.1 原理概述



Token (词元) 基础概念与作用

定义：文本拆分后的有意义片段，可为单字符、单词、词组

处理逻辑：文本转 Token 再映射为向量，实现计算机可读运算

生成逻辑：模型逐 Token 预测输出，依托输入与已有上下文生成完整文本

Token 的核心衡量 与商用计费价值

量化作用：Token 数量代表文本规模、计算量，数量越大算力需求越高

评测指标：Token 处理速度、生成准确率是大模型核心性能评判标准

收费规则：商用大模型按输入 / 输出 Token 计费，国内大模型资费远低于海外同类产品

大模型底层技术根基

技术底座：依托深度学习，海量数据、算力训练海量参数神经网络

基础架构：核心采用 Transformer 编码 - 解码架构，适配自然语言处理

Transformer 架构 与大模型核心能力原理

关键机制：依靠自注意力机制计算词与词关联权重

工作流程：文本 Token 转为向量传入网络，叠加词语相关性编码

最终效果：实现自然语言理解、文本生成，具备基础逻辑推理能力



1.3.1 原理概述





1.3.2 大模型训练的实例演示



未来学生 在校学生 教职员工 厦大校友 公共访问 我要捐赠 | English

首页 厦大概况 学部院系 人才培养 学术科研 社会服务 大学文化 合作交流 招生就业 人才招聘 学习二十大精神

厦大概况

厦门大学（Xiamen University），简称厦大（XMU），由著名爱国华侨领袖陈嘉庚先生于1921年创办，是中国近代教育史上第一所华侨创办的大学，也是国家“211工程”和“985工程”重点建设的高水平大学。

厦大概况

学校简介

厦门大学章程

当前位置: 首页 >> 厦大概况 >> 学校简介

学校简介

厦门大学（Xiamen University），简称厦大（XMU），由著名爱国华侨领袖陈嘉庚先生于1921年创办，是中国近代教育史上第一所华侨创办的大学，也是一所与中国共产党同龄的大学。在建校100周年之际，中共中央总书记、国家主席、中央军委主席习近平向学校发来贺信。贺信指出，厦门大学是一所具有光荣传统的大学。100年来，学校秉持爱国华侨领袖陈嘉庚先生的立校志向，形成了“爱国、革命、自强、科学”的优良校风，打造了鲜明的办学特色，培养了大批优秀人才，为国家富强、人民幸福和中华文化海外传播作出了积极贡献。



1.3.2 大模型训练的实例演示



公开网页文本 可作为训练数据

全球可访问的公开网页文本属于公开数据集，会被大模型开发者收集，用于大模型训练学习

文本预处理Token 切分

原始文本无法直接输入模型，需先切割为最小语义单元 Token，可通俗理解为独立字词，是模型训练的基础预处理步骤

数据格式转化词向量化

大模型仅识别数值型数据，需通过词向量化技术，将每一个 Token 转化为多维向量，完成文本到数值的转换

向量空间原理

为方便人类理解，可借助一维、二维、三维向量举例说明空间映射逻辑；主流大模型实际将 Token 映射至1000 维以上高维空间，以此精准捕捉丰富的语义关联

训练数据形态

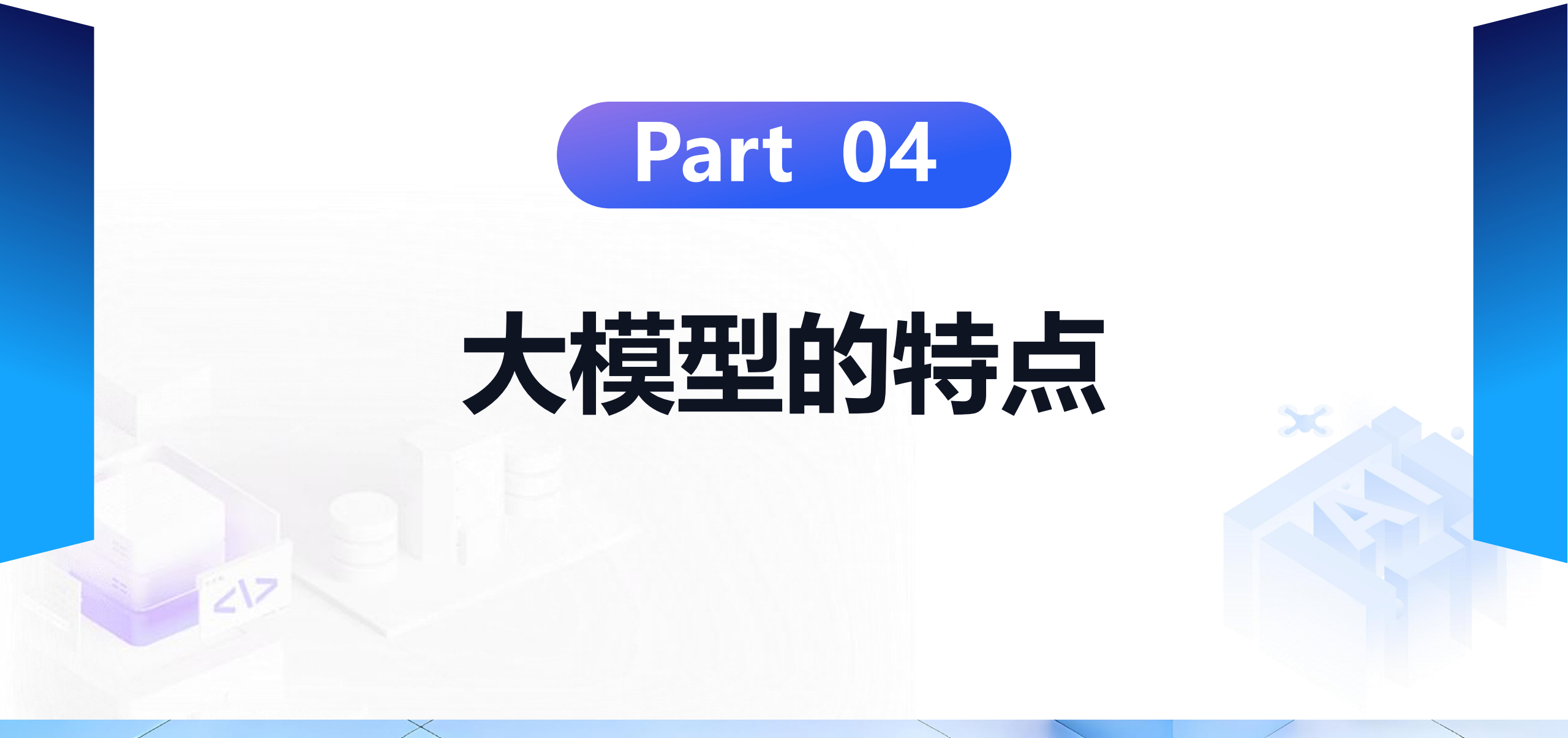
所有 Token 对应的高维向量整合后，形成数学矩阵，是输入大模型的核心数据形式

大模型训练的数学本质

大模型底层依托神经网络运行，整体训练过程，本质就是对文本向量矩阵进行一系列运算学习

Part 04

大模型的特点





1.4 大模型的特点



01 巨大的规模

02 涌现能力

03 更好的性能和泛化能力

04 多任务学习

05 大数据训练

06 强大的计算资源

07 迁移学习和预训练

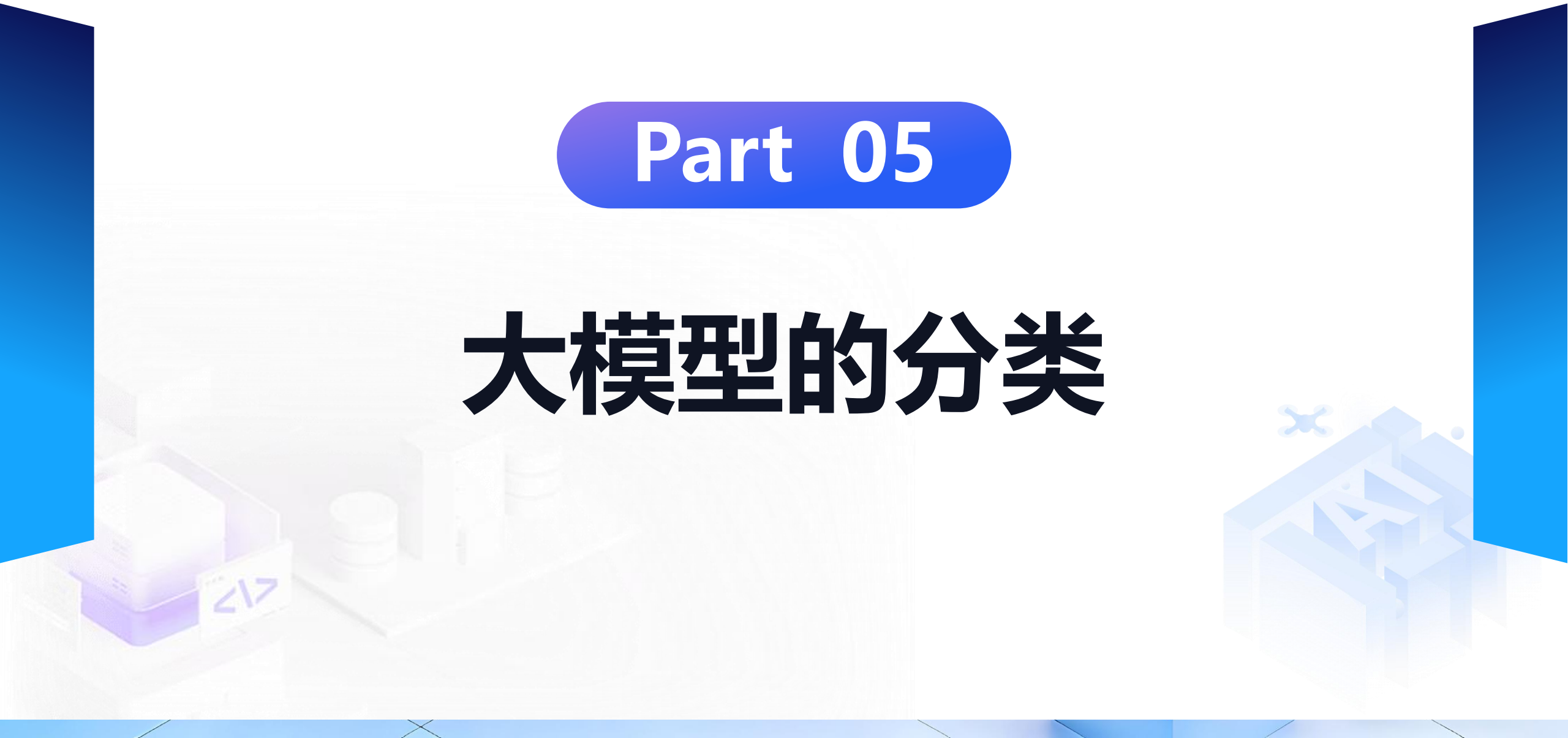
08 自监督学习

09 领域知识融合

10 自动化和效率

Part 05

大模型的分类





1.5 大模型的分类



按照输入数据
类型划分

按照应用领域
划分



1.5.1 按照输入数据类型划分



语言大模型

也称为“大语言模型”（LLM, Large Language Model），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

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图像分析。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

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



1.5.2 按照应用领域划分



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



1.5.3 大语言模型的分类

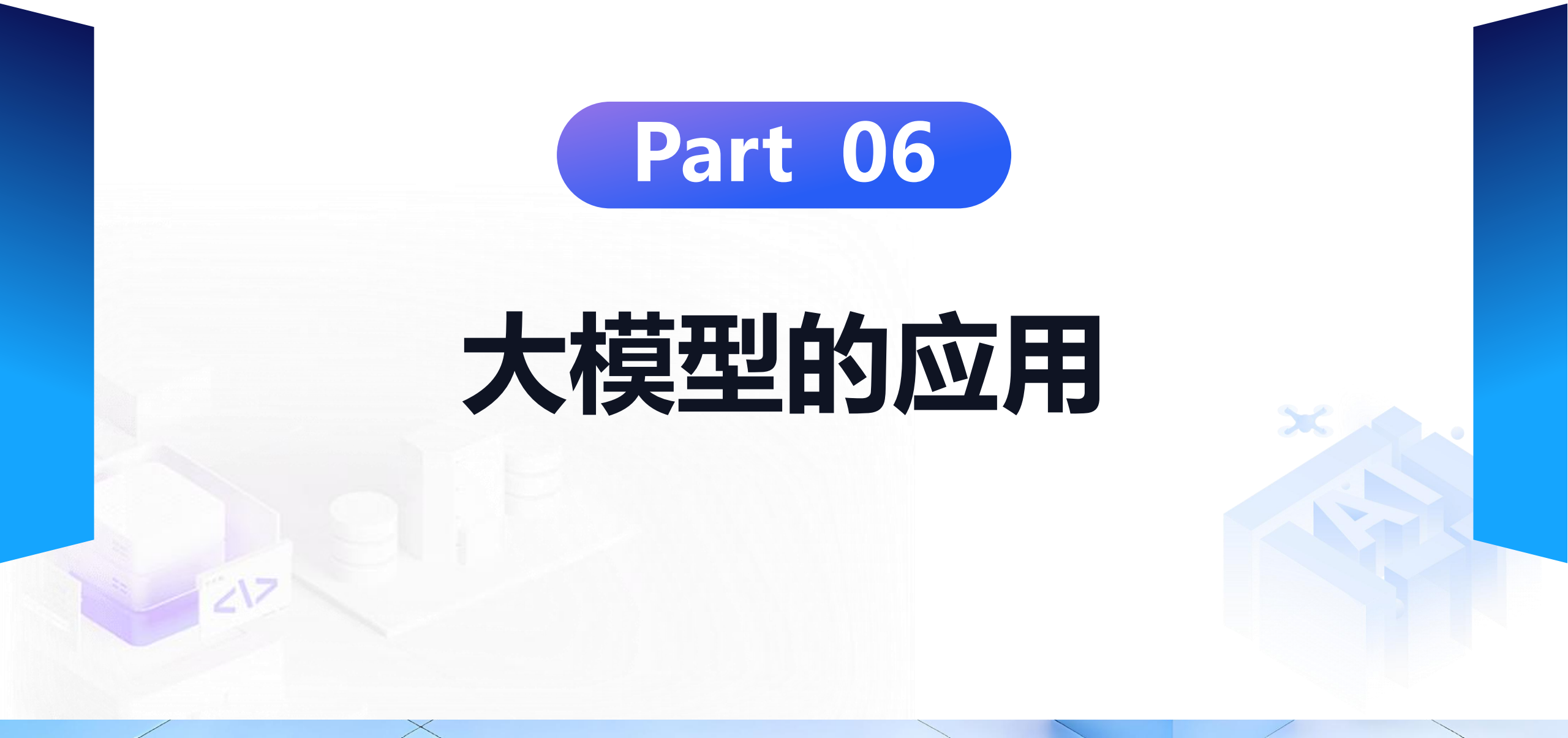


通用大模型和推理大模型的对比

特性	推理大模型	通用大模型
适用场景	复杂推理、解谜、数学、编码难题	文本生成、翻译、摘要、基础知识问答
复杂问题解决能力	优秀，能进行深度思考和逻辑推理	一般，难以处理多步骤的复杂问题
运算效率	较低，推理时间较长，资源消耗大	较高，响应速度快，资源消耗相对较小
幻觉风险	较高，可能出现“过度思考”导致的错误答案	较低，更依赖于已知的知识和模式
泛化能力	更强，能更好地适应新问题和未知场景	相对较弱，更依赖于训练数据
擅长任务举例	解决复杂逻辑谜题，编写复杂算法，数学证明	撰写新闻稿，翻译文章，生成产品描述，回答常识问题
成本	通常更高	通常更低

Part 06

大模型的应用



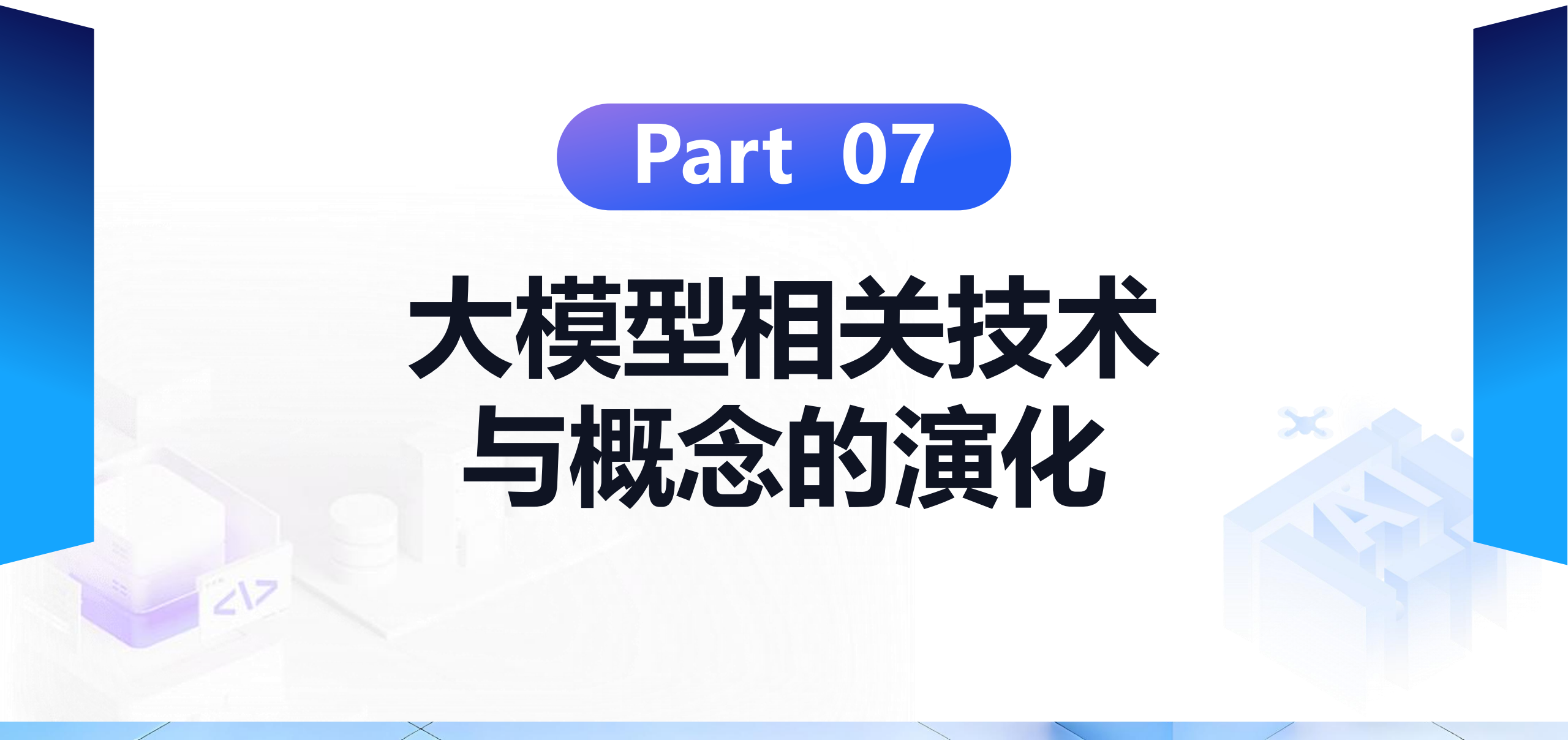


1.6 大模型的应用



Part 07

大模型相关技术与概念的演化





1.7 大模型相关技术与概念的演化



Token与
上下文窗口

提示词与
提示词工程

RAG与外部
知识增强

工具调用
与模型行动
能力

MCP与工具
生态连接

上下文工程
与
上下文组织

Skill与可复
用能力沉淀

Computer
Use与图形
界面操作

智能体与自
主任务执行

驾驭工程与
可靠性保障

工作流与业
务流程编排

工作空间智
能体与工作
空间智能化

技术演化的
主线



1.7.1 Token与上下文窗口



1 Token是理解大模型输入输出能力的重要基础

2 大模型推理成本和输入、输出 Token 数量相关

3 大模型单次推理可承载的信息量通过 Token 数量衡量

4 上下文窗口指模型单次推理可纳入计算的最大上下文范围，用户问题、历史对话、系统指令、检索文档片段、工具返回结果等内容共同占用该窗口

5 上下文窗口越大，模型可读取的素材越多，但不代表推理结果一定更优

6 上下文窗口过短会造成关键信息不足；无筛选的过长上下文会引入噪声，干扰模型抓取核心条件

7 大模型应用除关注模型可承载的信息总量外，还需明确信息供给范围、排序方式、压缩与过滤方法

8 信息筛选、排布、压缩过滤相关需求，衍生出提示词设计与上下文工程相关技术方向



1.7.2 提示词与提示词工程



01 提示词 (Prompt) 定义：用户或系统向大模型描述任务、背景、约束和输出要求的文本

03 提示词工程核心：不存在万能指令，核心是用清晰、完整、可执行的方式表达任务

05 面向计算机专业学生的通俗理解：提示词等同于面向大模型的任务说明书，描述越精确，模型输出越贴合预期

07 仅依靠提示词的局限性：处理内部文档、专有代码、训练数据后新增信息类任务时，复杂提示词也无法保障答案可靠

09 配套解决方案：大模型应用需引入外部知识检索机制

02 提示词工程诞生背景：相同任务使用不同提示词，大模型输出的质量、结构、可用性差异明显，由此诞生提示词工程实践方向

04 规范提示词包含要素：角色或任务定位、输入背景、目标要求、输出格式、约束条件、示例、评价标准

06 提示词工程能力边界：仅解决向模型交代任务的问题，无法弥补模型事实依据缺失的问题

08 模型缺失素材的后果：易生成逻辑通顺但不符合真实情况的内容



1.7.3 RAG与外部知识增强



RAG 全 称 是 Retrieval-Augmented Generation，译为“检索增强生成”，核心原理是模型生成答案前，先从外部知识库检索相关资料，结合用户问题与检索结果生成回答

相较于单纯依靠模型参数记忆，RAG 可降低过时知识、领域知识缺失、事实幻觉带来的风险

典型 RAG 系统包含文档处理、向量化表示、索引构建、检索排序、生成回答五大核心环节

RAG 预处理流程：文档入库前需切分片段，通过嵌入模型转化为向量，依托向量数据库实现向量的高效存储与检索；用户提问时，问题会同步转为向量，匹配知识库中语义相关的文档片段

RAG 实现了大模型从“仅凭参数记忆回答”到“结合外部资料回答”的转变，是大模型应用工程化的重要进展

基础 RAG 存在局限性，仅完成单次检索与单次生成，处理复杂问题时易出现检索不充分、问题拆解不足、资料冲突未处理等问题

Agentic RAG 可弥补基础 RAG 短板，支持系统依据中间结果改写查询、补充检索、拆分子问题、交叉验证来源，适配复杂研究型任务



1.7.4 工具调用与模型行动能力



RAG 仅提升大模型文本知识获取能力，模型输出形式仍以文本建议为主，Tool Calling/Function Calling 机制可推动模型从“提出建议”升级为“执行动作”，该机制支持模型按需选用外部工具并按预设参数格式发起调用



工具调用可对接搜索服务、数据库、代码环境、日历、订单系统、企业内部接口等外部能力；以查询订单状态为例，系统会先调用订单查询接口，再依据返回结果生成回答，模型完整参与理解意图、选择工具、填充参数、解释结果全执行链路



工具调用落地衍生多项工程难题：需定义各工具功能、参数、返回值与错误处理逻辑，同时做好权限管控、调用日志记录、危险操作限制、工具调用失败处理；工具数量增多后，若无统一连接方案，集成与维护成本会快速上涨



工具调用集成维护成本攀升的行业背景，推动标准化工具连接协议 MCP 诞生



1.7.5 MCP与工具生态连接



MCP (Model Context Protocol) 是面向模型应用的上下文和工具连接协议，目标是以相对统一方式向大模型应用暴露外部数据源、工具与资源

01

MCP 核心侧重点并非提升模型参数自身能力，而是降低大模型应用对接外部系统的复杂程度

02

无统一协议时，各类 AI 应用需单独适配文件系统、数据库、代码仓库、浏览器、企业系统等工具，不同工具在描述形式、参数传递、权限边界、返回格式上均存在差异

03

MCP 搭建统一连接层，让模型应用可采用标准化方式完成工具发现、资源读取、能力调用、结果接收全流程操作

04

从工程视角，MCP 体现大模型应用从单一聊天界面面向完整工具生态演进的行业趋势

05

大模型的综合价值不只由文本生成能力决定，同时取决于其能否规范、稳定、安全地对接现实工作场景的数据与业务系统

06



1.7.6 上下文工程与上下文组织



大模型应用发展趋势：从简单问答转向复杂 workflow，系统核心问题从处理单条提示词转变为整体上下文组织



上下文工程与提示词工程核心区分：提示词工程聚焦指令表述方式，上下文工程聚焦模型任务执行过程中可见信息范围



上下文工程核心目标：不单纯扩充输入长度，提升上下文的相关性、完整性、安全性



复杂 AI 系统核心认知：上下文组织的质量，重要程度与模型自身能力相当



上下文工程定义：模型调用前，系统针对性完成相关信息的选择、过滤、排序、压缩、组装，为模型提供适配任务的对应信息



多信息混合输入存在弊端：以智能客服为例，任务所需多类信息全部直接置入上下文，冗余、过期、无权限信息会干扰模型输出结果



上下文工程的系统定位：属于上下文组织系统，运行时自动完成资料检索、历史记录筛选、长文本压缩、敏感内容过滤、关键信息合理排布工作



1.7.7 Skill与可复用能力沉淀



Skill (技能) 定义

面向某一类任务的可复用能力描述；对比一次性提示词，Skill 侧重沉淀稳定任务流程、输出格式、质量标准、工具使用方式，支持同类任务重复使用既定方法



适用场景痛点与 Skill 解决价值

撰写周报、审查代码、分析数据表、生成课程讲义等步骤与标准固定的任务，若每次重写完整提示词会低效、易遗漏；Skill 可沉淀个人工作习惯、固化团队标准流程、转化组织岗位经验为可复用 AI 能力



Skill 核心意义

推动大模型应用从临时对话转变为可积累、可迁移的工作方法；不直接替代模型本身能力，而是为模型提供稳定的任务执行范式



Skill 拓展应用价值

Skill 结合工具调用、上下文工程后，AI 系统处理高重复、流程明确的专业任务的效果能够得到提升



1.7.8 Computer Use与图形界面操作



工具调用

工具调用大多依靠外部系统 API 接口，但实际工作存在不少无规范接口、接口成本高、权限复杂、仅支持人工图形界面操作的系统

Computer Use 定义

AI 系统观察屏幕、解析界面结构，通过鼠标键盘模拟人工操作计算机

Computer Use 典型任务

打开网页、识别按钮、填写表单、上传文件、下载报表、选择下拉菜单、复制粘贴内容、多页面切换

Computer Use 价值

拓展 AI 行动范围，可处理未做 API 改造的老旧系统、网页后台

Computer Use 存在稳定性、安全层面的多项难题

界面布局易变动、按钮识别易出错、登录验证码会阻断流程，单步误操作会干扰后续全部执行结果

Computer Use 适用场景与配套管控要求

仅适合受控环境作为辅助执行能力，需配套权限限制、操作确认、日志记录、人工监督机制



1.7.9 智能体与自主任务执行



智能体 (Agent) 定义

是可围绕目标开展计划、行动、观察、调整的 AI 系统；区别于普通聊天机器人，不只单轮应答，可拆解目标为多步骤、选用工具执行、依据结果修正计划

智能体通用工作循环

理解目标→制定计划→执行动作→观察反馈→更新状态→推进后续步骤

编程特性

编程是智能体落地实用价值较早的场景，依托代码项目具备清晰文件结构、可运行测试、可追溯报错、版本管控的特性；编程智能体可读取项目、定位函数、修改代码、运行测试，并依据测试失败结果迭代优化

编程智能体的发展

AI IDE、AI 编程智能体的发展印证智能体发展趋势，同时改变程序员工作模式：开发者除逐行编码外，还需侧重需求表达、架构设计、任务拆解、代码审查与质量管控

智能体能力提升弊端

智能体能力提升的同时需设定清晰边界，若无约束会出现错误代码修改、生成不安全代码、运行成本过高、执行流程无法追溯等问题



1.7.10 驾驭工程与可靠性保障



驾驭工程 (Harness Engineering) 定义

围绕智能体搭建的控制、验证、安全保障体系；模型为执行核心，Harness 包含权限控制、工具白名单、执行沙箱、日志追踪、错误重试、输出验证、人工审批、成本控制、回滚机制各类外围工程设施

智能体在不同环境核心需求存在差异

演示环境看重智能体自主完成任务的展示效果；生产环境核心要求是智能体可被控制、验证、审计

生产场景风险实例

文件整理智能体不可随意删除重要目录，代码修改智能体不可直接改动生产分支，业务审批智能体不可越权处理敏感事务；驾驭工程的核心作用是为智能体能力划定运行边界

企业真实需求

并非无限制自由的智能体，而是具备安全、可控、可追踪、可验证特性的智能体

大模型落地业务的工程难点转变

重心从评判模型智能程度，转移为保障整套智能体系统运行可靠性



1.7.11 工作流与业务流程编排



1

真实业务特征

真实业务具备多步骤、多系统、多角色的流程特征，无法依靠单次问答或单次工具调用完成

2

工作流的定义

将 AI 能力、API、数据库、消息系统、定时任务、条件判断、人工审批等环节按业务逻辑编排，保障任务持续、稳定运行

3

AI承担职责

工作流中 AI 仅负责部分判断、生成、分类任务，其余步骤由确定性程序、规则系统或人工节点承担

4

工作流与智能体

工作流与智能体不存在互斥关系，二者侧重点不同：智能体侧重面向目标的动态决策，工作流侧重业务步骤的稳定编排

5

实际业务系统

实际业务系统通常结合工作流与智能体使用：工作流负责流程可管理、可复现，智能体处理流程内需灵活判断、信息综合的环节



1.7.12 工作空间智能体与工作空间智能化



工作空间智能体

(Workspace Agent) 是面向团队或组织工作空间的长期智能体形态，区别于一次性任务智能体，工作空间智能体长期驻留于项目、文档、任务、会议、代码仓库、协作系统内，可理解团队上下文，持续承担对应岗位辅助工作

01

工作空间智能体可调用项目进度、角色分工、历史文档、会议纪要、任务状态、客户记录、权限规则等多类信息提供服务

02

工作空间智能体代表 AI 产品形态迭代升级：AI 从临时使用工具转变为组织工作空间内的固定数字角色，涵盖研发、销售、项目管理、数据分析等各类助理角色

03

工作空间智能体衍生出更复杂的治理相关问题，涉及权限管理、数据隔离、责任归属、记忆更新、自动执行边界、人工确认机制多个维度

04



1.7.13 技术演化的主线



大模型相关技术概念并非彼此孤立，而是其应用贴近真实工作场景过程中形成的完整技术谱系，各概念对应不同工程问题

Token 与上下文窗口

界定模型单次可处理的信息体量

提示词与上下文工程

规范向模型传递任务与材料的方式

RAG 与 Agentic RAG

实现模型接入外部知识的需求

工具调用与 Computer Use

支撑模型完成各类执行动作

MCP 与驾驭工程

实现标准化连接，保障模型应用可靠性

Skill、工作流与 工作空间智能体

实现技术能力沉淀，适配组织协作场景



1.7.13 技术演化的主线



大模型整体发展趋势：
从单一聊天工具，迭代
为可参与实际工作的系
统组件



**大模型能力提升覆盖内
外双维度，不仅包含模
型内部能力升级，还涵
盖数据组织、工具连接、
权限控制、流程编排、
质量评估、协作治理等
外部配套能力**



**计算机类专业学生学习
大模型技术的核心：无
需机械记忆术语定义，
重点把握各类技术对应
的工程解决问题**



**大模型技术学习需遵循
固定演化主线：信息表
示 — 任务表达 — 知识
增强 — 工具执行 — 系
统可靠性 — 组织协作**





1.7.13 技术演化的主线



大模型时代相关技术概念的层次关系

层次	代表概念	主要问题
信息层	Token、Context Window	模型一次能够处理多少信息
指令层	Prompt、Context Engineering	如何表达任务并组织上下文
知识层	RAG、Agentic RAG	如何接入外部资料并降低幻觉
执行层	Tool Calling、Computer Use	如何从生成文本发展为执行动作
工程层	MCP、Harness Engineering	如何标准化连接、控制权限并保证可靠性
组织层	Skill、Workflow、Workspace Agent	如何沉淀能力并进入团队协作



1.8 本章小结



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！

林子雨 副教授 厦门大学

附录A：主讲教师林子雨简介

主讲教师：林子雨

单位：厦门大学计算机科学与技术系

E-mail: ziyulin@xmu.edu.cn

个人网页: <http://dblab.xmu.edu.cn/post/linziyu>

数据库实验室网站: <http://dblab.xmu.edu.cn>



- 林子雨，男，1978年出生，博士（毕业于北京大学），全国高校知名大数据与人工智能教师，入选“2021年高校计算机专业优秀教师奖励计划”。现为厦门大学计算机科学与技术系副教授，厦门大学信息学院实验教学中心主任，曾任厦门大学信息科学与技术学院院长助理、晋江市发展和改革局副局长。中国计算机学会数据库专业委员会执行委员，中国计算机学会信息系统专业委员会执行委员。
- 国内高校**首个“数字教师”提出者和建设者**，厦门大学数据库实验室负责人，厦门大学云计算与大数据研究中心主要建设者和骨干成员，2013年度、2017年度、2020年度、2023年度和2026年度厦门大学教学类奖教金获得者，荣获2024年福建省高等教育教学成果奖特等奖（个人排名第七）、2022年福建省高等教育教学成果奖特等奖（个人排名第一）、2018年福建省高等教育教学成果奖二等奖（个人排名第一）、2018年国家精品在线开放课程、2021年国家级线上一流本科课程、2020年国家级线上一流本科课程。**编著出版了20余本大数据与人工智能系列教材，被国内1000多所高校采用，3本教材入选教育部“十四五”普通高等教育本科国家级规划教材**；建设了**国内高校首个大数据课程公共服务平台**，为教师教学和学生学习大数据课程提供全方位、一站式服务，年访问量超过400万次，累计访问量超过3200万次。**大数据系列MOOC课程入选“2023年教育部国家智慧教育公共服务平台应用典型案例”**。2025年发布4个大模型科普报告，全网累计浏览量超过1000万，并应邀为高校、政府和企业做170余场大模型讲座。
- 主要研究方向为数据库、数据仓库、数据挖掘、大数据、云计算和物联网，并以**第一作者身份在《软件学报》《计算机学报》和《计算机研究与发展》等国家重点期刊以及国际学术会议上发表多篇学术论文**。作为项目负责人主持的科研项目包括1项国家自然科学基金青年基金项目(No.61303004)、1项福建省自然科学基金青年基金项目(No.2013J05099)和1项中央高校基本科研业务费项目(No.2011121049)，主持的教改课题包括1项2016年福建省教改课题、1项2016年教育部产学协作育人项目、1项2024年教育部产学协作育人项目。

附录B：大数据学习路线图



大数据学习路线图访问地址: <http://dblab.xmu.edu.cn/post/10164/>

附录C：林子雨大数据系列教材



了解全部教材信息: <http://dblab.xmu.edu.cn/post/bigdatabook/>

附录D：林子雨人工智能通识系列教材

厦门大学林子雨**人工智能通识**系列教材
被众多高校采用，满足不同高校差异化教学需求



了解全部教材信息：<http://dblab.xmu.edu.cn/post/ai-book/>

附录E：林子雨Python系列教材



了解全部教材信息: <https://dblab.xmu.edu.cn/post/python-books/>

谢谢!



林子雨 副教授

厦门大学

