

《人工智能与旅游创新应用》

生成式人工智能 功能解析

模块一 生成式人工智能发展概况

旅游管理专业核心课程

主讲教师： ____ · 2026 秋季学期

MODULE 01 — PROJECT 03



CONTENTS · 本讲目录

从"会用"到"懂它"

PANORAMA TEXT
VISION · PRINCIPLE

01 功能全景

生成与判别 · 五大功能域 · 一日旅程触点 · 规模数据

02 文本能力

理解生成推理 · 旅游文本场景 · 适配与陷阱

03 视觉·语音·多模态

图像 · 场景图谱 · 语音 · 视频 · 多模态融合

04 原理与边界

演化脉络 · 预测机制 · 训练三部曲 · 幻觉与边界

学习目标

三维目标

知其然 知其所以然

KNOW WHAT
KNOW WHY

知识目标

能列举生成式 AI 的**五大功能域**及各自成熟度；能用自已的话说清 " 预测下一个词 " 的基本机制。

能力目标

能判断一个旅游任务**适合交给 AI 的哪种能力**，能识别生成内容的幻觉与套路化痕迹。

素养目标

建立 "**能力有边界**" 的技术观：不神化也不矮化 AI ，做清醒的使用者。

一个词，两种 AI

过去十年你用过的 AI 都在 " 判别 "—— 刷脸进门、地图导航、OTA 推荐酒店，AI 在答案里挑。今天的大模型在 " 生成 "—— 你让它写一段西塘夜游文案，它从无到有地造。从挑到造，一字之差：判别式在已知里选择，生成式在未知里创造。

" 生成 " 为什么比 " 判别 " 难得多？

AI 的 " 造 " 与人类的 " 创作 "，是一回事吗？

它能造什么？边界在哪里？

生成

GENERATE

01

PART

功能 全景

THE CAPABILITY P/

① 生成与判另 ② 五大功能域 ③ 一日旅程 ④ 规模数据

生成式 AI：学的是 " 分布 "，不是 " 规则 "

判别式 AI 在已知答案里挑，**生成式 AI** 在无限可能里造——前者是阅卷老师，后者是考生。

—— 综合主流教材与业界通行表述

判别式

DISCRIMINATIVE

人脸识别、垃圾邮件过滤、酒店推荐——
先有人定好类别，AI 学习**分类的边界**，输出是 " 属于哪一类 "。

生成式

GENERATIVE

写文章、画海报、配音——AI 学习海量数据后掌握**内容的统计规律**，输出是 " 一个全新的样本 "。

大模型

LLM

生成式的当代主角：参数巨大、吃下全网文本——**ChatGPT 与国产五强**都属于这一家族。

分界线不在 " 智能高低 "，而在**输出形态**——一个是判断，一个是创造。

五大功能域：AI 的 " 手艺清单 "

01

文本

成熟度 ★★★★★

写攻略、改文案、译外文——**最成熟**，旅游人第一生产力。

02

图像

成熟度 ★★★★★

海报、插画、效果图——**以假乱真**，创意行业震动最大。

03

语音

成熟度 ★★★★★

配音、翻译、识别——讲解员的**数字嗓子**。

04

视频

成熟度 ★★★

文生视频、数字人口播——**进步最快**，可控性最低。

05

多模态

成熟度 ★★★

看图说话、图文互转——**融合趋势**，下一代交互入口。

本讲按 " 成熟度 " 顺序逐项拆解：先文本，后视觉，再展望融合——**先吃透最稳的，再观望最新的。**

一日旅程： AI 功能触点地图

07:00 · 行前

出发

- 行程推送 (文本)
- 天气播报 (语音)

11:00 · 游中

抵达

- 扫码识景 (图像)
- AR 导航 (多模态)

21:00 · 游后

分享

- Vlog 脚本 (文本)
- 旅拍配乐 (音频)

五大功能域不是实验室概念——你的一天早就被它们串起来了。这门课要做的，是把 " 被动被服务 " 变成 " 主动去驾驭 "。

功能爆发：三组数字

ChatGPT 发布后

2
个月

月活用户突破 1 亿——史上最快的消费级应用，生成能力第一次大规模交到普通人手上。

86%

企业采用率

麦肯锡 2024 年调研：超八成企业已在至少一个业务环节使用生成式 AI，营销与客服**渗透最深**。

5 种

模态覆盖

从纯文本起步，两年内扩展到图像、语音、视频、代码——**能力的边界每个月都在外扩**。

本期

旅游业的反应

携程、同程、马蜂窝、飞猪相继上线 AI 助手——**不是要不要用，是怎么用好**。

02

PART 1

文本能力

TEXT GENER.

① 两翼 ② 三支柱 ③ 旅游场景 ④ 适配与陷阱

文本能力两翼：写得像，懂得多

翼一 · FLUENCY

写得像

语法通顺、辞藻到位、结构工整——**表层语言能力**。像一位中文系毕业生：你出题，它成文，口吻随调。

翼二 · KNOWLEDGE

懂得多

训练时 " 读 " 过的知识横跨百科、论文、论坛——**知识广度**。像一位杂学通才：天文地理都能聊，深度另说。

两翼缺一不可：只 " 像 " 会空洞，只 " 懂 " 会干瘪。用得越久越能体会——**提示词是在调度这两翼的配比**。

三支柱：理解 · 生成 · 推理

支柱① · 理解

读得懂

输入端能力

- 归纳 500 字差评的**三个不满**
- 从行程单算出**空闲时段**
- 长攻略提炼**要点清单**

支柱② · 生成

写得出

输出端能力

- 小红书风 / 公文风**一键切换**
- 按指定字数结构**批量产出**
- 同素材**多版本**改写

支柱③ · 推理

想得通

思考端能力

- **条件规划**：下雨则行程 B 替代 A
- 费用**多方案**测算
- 投诉处理**推演**

强度规律 生成最稳，理解次之，**推理最容易翻车**——多步复杂推演务必人工复核。

旅游文本： AI 最肥沃的试验田

结构模板化

行程 = 天数 × 景点 × 交通 × 餐饮，**天然适合 AI 填充框架**——这是旅游文本区别于小说、诗歌的关键。

四大高频场景

行程规划**初稿**十秒出；攻略文案**量产**；导游词**分风格改写**（学术版 / 亲子版 / 相声版）；客服话术**模板库**即时生成。

为什么最适配

信息密度低、情感浓度高——AI 的 " 套路化表达 " 在种草文案里**反而成立**，读者要的就是那个味道。

一个判断标准

这活儿有没有**固定框架**？

要不要**查实时事实**？

框架越固定，AI 越顺手；

事实越实时，你越要盯。

一句话：旅游内容是 " 戴着镣铐跳舞 "——而**镣铐**，恰恰是 AI 的舒适区。

任务适配：什么活交给它，什么活留给自己

文本任务	AI 表现	人的角色
批量初稿（攻略 / 文案）	优——快且稳	定方向、给素材
长文改写（换风格 / 换长度）	优——一遍到位	验收口吻是否对味
信息归纳（评论 / 报表总结）	良——偶有遗漏	抽查关键数据
多约束行程规划	中——细节易错	核验班次、营业时间
事实性查询（门票 / 政策）	差——幻觉高发	必须自查官网

规律：越 " 创作型 " 越放心交给 AI ，越 " 事实型 " 越要自己把关——**创作它管，事实你管。**

三个坑：写得越好，越要警惕

坑一 幻觉：一本正经地编造

编造不存在的 " 网红餐厅 "、过期班次、" 查无此地 " 的小众秘境——**语气越自信，越要核实。**

坑二 套路化：千文一面

" 绝绝子 "" 宝藏 "" 秘境 " 批量复读——读者审美疲劳，**平台限流**，AI 味成了负资产。

坑三 文化误读：懂词不懂礼

把民族习俗、宗教禁忌当普通素材混排——旅游文本里，**文化错误是红线**，AI 没有敬畏心。

对策

事实→ 关键信息三源核验

文风→ 人工润色去 AI 味

文化→ 敏感内容专家审读

课堂辩论：AI 导游词有没有 " 灵魂 "

某景区计划用 AI 批量生成导游词，讲解员小组对此发生分歧——一场关于 " 讲解本质 " 的争论开始了。

—— 议题背景

正方 · 有灵魂

- 讲解的本质是**信息传递**，AI 信息量更足、更准
- 游客要的是 " 听懂 " ，不是讲解员的个人风采
- AI 能按人群**实时换版本**：亲子版、老年版——人力做不到

VS

反方 · 没灵魂

- 讲解是**情感劳动**：眼神、停顿、共情，AI 没有身体经验
- 导游词的 " 魂 " 来自**在地生活**——AI 只有语料，没有生活
- 当所有景区用同一批 AI ，**千景一面**，体验加速贬值

讨论规则：每方 **2 分钟立论** + 自由交锋。**辩完做一件事：给 " 灵魂 " 下个可检验的定义。**

03

PART II

视觉·语音 多模态

VISION · VOICE · MU

① 图像 ② 场景图谱 ③ 语音 ④ 视频 ⑤ 融合

图像生成：文生图与图生图

路线一 · TXT2IMG

文生图

一句提示词 → 一张图。**想象空间大**，可控性低。例："水墨风格的黄山云海，留白，宣纸质感"——适合概念稿、氛围图。

路线二 · IMG2IMG

图生图

一张参考图 + 指令 → 改风格、换背景、扩画面。**可控性高**。例：把实拍照片转成手绘插画风，保留构图——适合批量改版、风格统一。

旅游实操口诀：**概念用文生图，落地用图生图**——先发散，再收敛。

旅游视觉：四类高频应用

视觉能力

从海报到数字人

视觉是旅游营销的**第一入口**——用户先“看图”，再决定“读文”。

图像生成已渗透进内容生产的**每一个环节**，从宣发到体验。

01

宣发海报

节庆、活动、线路海报**快速出稿**，一版三稿比价。

02

社媒配图

小红书封面、公众号头图**量产**，风格统一。

03

场景概念

未建成的营地、民宿改造**效果图先行**，低成本预演。

04

讲解插图

历史场景**复原想象图**：唐长安西市、宋清明上河图细节。

语音：被低估的一环

三项基本功能

合成 TTS：文本转语音，数字嗓子——讲解配音、有声攻略；**识别 ASR**：语音转文本，数字耳朵——会议纪要、留言整理；**实时翻译**：数字译员——入境接待即时沟通。

旅游场景红利

旅游服务**高度依赖 " 说 " 与 " 听 "**：讲解、接待、客服——语音 AI 直接对接这些环节；方言与多语种接待长期缺人，语音翻译正在补位。

现状提醒

合成语音的**情感颗粒度**仍在爬坡——正式场合的深情讲解、高端接待，真人配音仍不可替代。

一分钟体验

随便打开一款大模型 App

找到**语音对话**开关

用方言问一个旅游问题

感受 " 耳朵 + 嗓子 " 的进化

语音是从 " 工具 " 走向 " 陪伴 " 的入口——**数字人、智能客服**都长在它上面。

视频：进步最快，最不可控

现状 · 能做到

秒级出片

文生视频 **5-10 秒短片**，可用于氛围渲染；图生视频：一张景区照片 → **推拉摇移运镜**的动态画面。代表工具：可灵、即梦、Sora。

短板 · STILL HARD

仍做不到

长叙事：超过十几秒情节一致性崩坏；**精确控制**：指定 " 第三秒转向牌坊 " 办不到；**文字画面**：店招、标语常成乱码。

旅游应用姿势：**AI 出空镜素材 + 真实拍摄主体 + 人工剪辑**——各干各的活。

案例：一张节气海报的诞生

CASE · 实操复盘

需求：民宿“谷雨茶山体验”节气海报，24小时内交付

传统流程 摄影约拍 2 天 + 设计排期 3 天 + 改稿 2 轮 ≈ 5 天

AI 流程 文生图 12 张初稿 + 图生图细化 + 人工排版 ≈ 3 小时

- **第一步 · 发散**：文生图出 12 张概念稿——雨雾、茶山、竹篮、屋檐四个意象各三张。
- **第二步 · 收敛**：选中“雨雾茶山”，图生图统一为水彩风，补一版竖版构图。
- **第三步 · 落地**：留出文字区，人工排版店名与二维码——**AI 管画面，人管信息。**

结果

3 小时交付，成本约为传统的**十分之一**——甲方选中的正是 AI 初稿的方向。

启示

AI 加速的是“**从 0 到 60 分**”；“从 60 到 90 分”的审美判断与信息把关，仍然是人的护城河。

多模态：能力开始“合体”

01 单模态时代 2022—2023

文本、图像、语音各自为战——换个模态要换工具，像带着一串遥控器出门。

02 多模态理解 2024

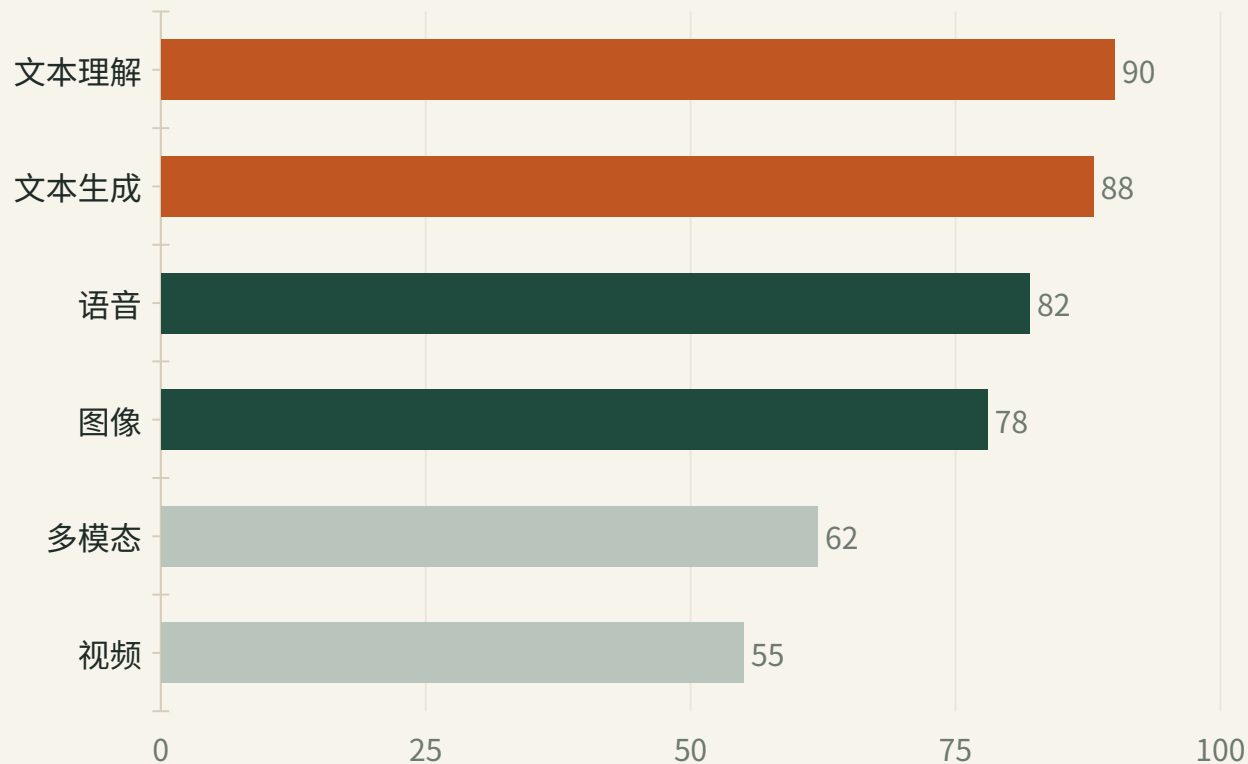
一个模型**看图能聊**：拍下菜单问“哪个辣”、传行程截图让它优化——“眼睛”长上了。

03 任意模态生成 2025—

输入输出自由组合：图文互相生成、语音驱动口播数字人——**工具开始消失，能力开始聚合。**

对你的意味 学工具的速度永远追不上工具迭代——**抓能力，不抓工具**：理解“它能干什么”，比记住“它叫什么”重要十倍。

能力成熟度：课堂示意图



课堂示意分数（满分 100），综合公开评测与实务感受，仅作参考。

01

文本双项最成熟——写作类任务可放心交。

02

图像、语音第二梯队——批量出稿可用，精品仍需人。

03

视频最年轻——正确姿势是“要素材，不要成片”。

04

PART F

原理 与边界

HOW IT WORKS · WHE

① 演化脉络 ② 核心机制 ③ 训练三部曲 ④ 幻觉与边界

从判别到生成：十二年长跑



今天的爆发不是奇迹，是**十二年渐进积累**的水到渠成。

如果只记一个年份，记 2017

2017

注意力机制元年

TRANSFORMER

08 页
纸

一篇论文

《Attention Is All You Need》——谷歌八位研究员的一篇技术报告。

新机
制

会 " 划重点 " 的 AI

读句子时**自动判断 " 谁和谁相关 "**——" 它 " 指代哪个景点，一眼看穿。

大地位

所有大模型的地基

GPT、文心、DeepSeek——**全部建在它之上。**

给 AI 装上了 " 划重点 " 的本事——从此带着理解读书，不再逐字。

核心机制：超级文字接龙

机制 · NEXT TOKEN

预测下一个词

给定前文，模型计算词表中**每个词的概率**，选一个接上，再继续。"苏州园林甲天下，"→ 下一字大概率 "江"→"南"——一个字一个字地 "接" 出整篇文章。

澄清 · NOT MAGIC

它没有在 "思考"

没有查数据库——答案是从**概率分布里采样**出来的；没有 "知道" 对错——**流畅≠正确**，它只负责像，不负责真。

记住这句话：**AI 是概率的艺术家，不是事实的守门员**——真假要靠你来把关。

大模型是怎么炼成的

三步走：博览群书 → 专科深造 → 学会做人——一步都不能省。

STEP 1 预训练 · 博览群书

吞下几乎整个互联网文本——只学“语言像不像”，不问真假对错。

STEP 2 微调 · 专科深造

用高质量问答数据教它按助手的方式对话——旅游、法律、医疗各有专才版。

STEP 3 对齐 · 学会做人

人类反馈打分，教它有用、诚实、无害——拒答危险问题、承认不知道。

一个类比

像培养一位导游

预训练像孩子读完整个图书馆；微调像大学选了旅游管理专业。

对齐像入职培训——知道什么该说、怎么说得体。

三个关键词，看懂产品发布会

关键词① · TOKEN

文字积木

计费单位

- 一个汉字 $\approx$ 1-2 个 token
- 对话按 token 计费——**长文 = 贵**
- 看报价先看 token 单价

关键词② · 参数量

脑容量

模型规模

- 7B / 70B / 万亿级
- 参数越多通常能力越强
- **但大 $\neq$ 一定好**，训练质量同样关键

关键词③ · 上下文窗口

工作记忆

记忆长度

- 一次能 " 记住 " 的对话长度
- 128K $\approx$ 一部中篇小说
- **超出的部分彻底忘记**

实用价值 下次看发布会，这三个词不再是天书——**都是可以追问的具体指标。**

边界：它天生做不到的三件事

边界一 不保证事实

概率生成机制决定它**没有 " 查证 " 动作**——门票价格、开放时间、政策条文，都可能过时或编造。

边界二 没有真实经验

它 " 读过 " 十万篇游记，但**没挤过一次早高峰的索道**——身体经验、在地判断，永远是人的领域。

边界三 不了解你的今天

训练数据有截止日，且**看不到你的私有信息**——今天的空房、你客户的偏好，要么不给它，要么告诉它。

正确姿势

事实→ 交给官方渠道

经验→ 人来补充在地判断

信息→ 把上下文喂给它

幻觉实拍： AI 推荐的 " 秘境 "

CASE · 事故复盘

情境：让 AI 推荐 " 杭州小众古村落，一小时可达 "

AI 回答 推荐 " 云栖坞古村 "：明清建筑、无门票、公交直达

实地核查 查无此村——地名真实存在，但是片茶山，无古建筑

- **幻觉机理**："云栖"（真实地名）+"坞"（常见后缀）+"古村"（你的要求）——**概率把它们粘成了看似合理的答案。**
- **识破要领**：越是 " 完美贴合需求 " 的小众推荐，越要警惕——**热门信息它学得多，冷门信息全靠编。**
- **处置流程**：地图核位置 → 官方号核开放状态 → **三源交叉**后才可写进行程。

行业警钟

2023 年起多家旅行社因 AI 生成行程中的**虚构景点**遭游客投诉——"AI 说的 " 不能成为免责理由。

职业底线

旅游从业者对事实的**核验责任**，不因使用 AI 而**转移**——工具升级，责任不降级。

边界速查：能做 · 不能做 · 需把关

功能域	可以放心做	必须人工把关
文本生成	初稿、改写、翻译、总结	全部事实类信息
图像生成	概念图、配图、风格化	涉及真人肖像与历史场景
语音合成	草稿配音、内部审听	对外发布的声音授权
视频生成	空镜、转场、氛围素材	成片叙事与文字画面
多模态	图文互转、截图理解	跨模态信息的一致性

把这张表贴在工位上：**每一类产出，都对应一道人工工序**——省掉工序，就是在积累风险。

课堂辩论：AI 越强，人越贬值？

学完功能与边界，回到那个绕不开的问题：这些能力最终会替代旅游人吗？两种立场，都有道理。

—— 议题背景

正方 · 会贬值

- 文案、设计、讲解的**基础产能**已严重过剩——初级岗位首当其冲
- 学习门槛降低，**会用 AI 的人**与不会用的人产出差十倍
- 客户也在用 AI——**信息不对称的红利**正在消失

反方 · 不会

- 旅游是**体验经济**：温度、在地、信任，机器给不了
- AI 拉平的是 "60 分"，**90 分以上的判断力反而更稀缺**
- 新技术消灭岗位，也**创造岗位**：AI 内容运营——十年前不存在

VS

讨论规则：每方 **2 分钟立论** + 自由交锋。**课后思考：你的护城河建在 60 分之上，还是 90 分之上？**

驾驭之道：三个使用原则

01

人机分工

DIVISION

它管快，你管对

创作交给 AI ，事实留给自己。每类任务先问：这件事的 " 真 " 由谁兜底？

02

验证习惯

VERIFY

像系安全带一样

关键信息三源交叉：官方渠道 + 地图 + 实地电话。把核验做成肌肉记忆。

03

场景匹配

CONTEXT

按观众定把关强度

内部草稿可以大胆用 AI ，**对外发布必须过**人——按 " 谁会看到 " 决定把关强度。

合起来一句话 把 AI 当能力极强的实习生用——活儿尽管派，署名与责任永远是你自己的。

带走的清单：从今天开始练

功能解析不是听懂，是用会的——本周完成以下七件事：

01 三口吻练习 让 AI 用学术 / 亲子 / 文艺三种口吻写同一段景点介绍

02 改写接力 官方新闻稿 → 小红书笔记 → 短视频脚本

03 图像实测 文生图、图生图各出 5 张，感受可控性差异

04 语音体检 语音合成录一段导游词，与真人版对比情感颗粒度

05 幻觉狩猎 故意让 AI 推荐小众景点，抓出至少一个幻觉

06 窗口实验 超长对话后问开头细节，亲测 " 工作记忆 " 的边界

07 边界建表 为自己岗位列一张 " 放心做 / 须把关 " 清单

四个问题，四个答案

Q1 生成式 AI 和判别式 AI 的本质区别？

输出形态：判别在已知里挑，生成在未知里造——学的是分布，不是规则。

Q2 五大功能域里，旅游人最该先吃透哪个？

文本——最成熟、场景最多、见效最快；图像其次。

Q3 为什么 AI 会 "一本正经地胡说八道"？

它按概率生成，只保证 "像"，不保证 "真"——预测下一个词的机制里没有查证环节。

Q4 一句话说清人机分工？

创作它管，事实你管；速度它管，责任你管。

作业：功能实测报告

思考题 · 下次课开篇讨论

- ① 本次实测中，哪个功能域的表现**最超出你预期**？哪个最让你失望？
- ② 结合你的目标岗位（导游 / 计调 / 营销 / 策划），**哪个能力**对你最有职业价值？

课后作业 · 实践任务

任务一

完成清单中的 " 幻觉狩猎 "：让 AI 推荐 5 个本地小众景点，逐个核验真实性，写 300 字复盘。
要求附对话截图与核验证据，标注哪个环节最容易翻车。

任务二

选定一个功能域，为家乡的一个景区产出一件 AI 作品（文案 / 海报 / 配音任选）。
同时提交 200 字 " 人机分工说明 "：哪些是你做的，哪些是 AI 做的。

提交：**下次课前**上传课程平台；优秀作品将进入课程案例库，署名展示。

延伸学习资源

论文

《Attention Is All You Need》（2017）

只读引言与图 1，感受 "划重点" 机制的原文表述——**八页纸改变了世界**。

报告

中国信通院《人工智能发展报告》（年度）

国内权威视角，**技术成熟度曲线**每年更新，看趋势必读。

纪录

生成式 AI 主题纪录片与公开课

各平台均有优质免费内容——**挑一部看透**，胜过刷十个短视频。

实操

主流工具的 "能力边界" 官方文档

每家官网都写了 "我们可能出错"——**读官方的自我坦白**，是最快的祛魅方式。

下讲预告前的小任务：**把 "预测下一个词" 讲给一位非计算机专业的朋友听**——讲得明白，才算真的懂。

项目四：生成式人工智能价值研判

MODULE 01 · PROJECT 04

能力已明， 价值几何

从 " 能做什么 " 到 " 该做什么 "

01 效率与体验

它到底省了多少成本，又**创造**了什么新价值？

02 就业与伦理

岗位会被替代吗？**数据与隐私**的红线在哪里？

03 研判框架

给自己一套**冷静评估** AI 价值的标尺，不吹不黑。