

《人工智能与旅游创新应用》

关注生成式 AI 工具 的应用挑战

模块二 生成式人工智能基础操作

旅游管理专业核心课程

主讲教师：_____ · 2026 秋季学期

MODULE 02 — PROJECT 04

CONTENTS · 本讲目录

会用， 更要慎用

RISK ·
HALLUCINATION
DATA ·
COMPLIANCE

01 风险图谱

挑战定义 · 事故数据 · 四大挑战 · 风险分级

02 幻觉防御

三副面孔 · 高发场景 · 核验五步 · 三道防线

03 数据安全

泄露代价 · 脱敏四法 · 分级策略 · 安全守则

04 合规生产

合规三层 · 内容标识 · 素材规范 · 发布清单

学习目标

三维目标

胆大心细

BOLD BUT
CAREFUL

知识目标

能列举生成式 AI 应用的**四大挑战**（幻觉 · 隐私 · 版权 · 依赖）及典型事故类型；说出国内监管对**内容标识**的基本要求。

能力目标

能执行**信息核验五步法**与**数据脱敏四法**；能按发布前检查清单完成一次合规自查。

素养目标

建立 "**责任不外包**" 的职业底线：工具升级，核验与把关的责任永远在使用者。

一条 AI 攻略，赔了八千块

旅行博主照搬 AI 生成的 " 小众玩法 " 发进攻略——
推荐的 " 悬崖咖啡馆 " 并不存在，粉丝按图索骥被
困山腰，救援加赔偿花了八千。评论区最扎心的一
条： " 博主自己没去过，就敢写攻略？ "

AI 错了，锅算谁的？

哪些信息必须核验，哪些可以放心？

有没有一套流程，让大胆用与不踩雷并存？

本讲交付： **核验五步 + 脱敏四法 + 发布清单。**



01

PART

风险 图谱

RISK LANDS

① 挑战定义 ② 事故数据 ③ 四大挑战 ④ 风险分级

应用挑战：工具好用，边界要清

应用挑战 = 使用生成式 AI 过程中**可能造成损失的风险点**——技术特性与使用方式共同决定风险大小。

——综合主流教材与业界通行表述

技术侧

TECH SIDE

幻觉（编造事实）、黑箱（无法解释）、时效（知识截止）——**机制的固有属性**，无法根除只能防御。

使用侧

HUMAN SIDE

隐私泄露、版权越界、过度依赖、权责不清——**取决于使用者行为**，规范操作即可大幅降低。

环境侧

CONTEXT

法规尚在完善、平台规则不一、判例持续出现——**合规成本在上升**，先行合规者反而获得信任优势。

关键认知：八成风险来自**使用侧**——也就是说，**八成风险可控**，这是本讲的立足点。

事故数据：三个切面

AI 相关投诉涉及事实错误

7 成

旅游平台客诉抽样：AI 生成内容引发的纠纷中，约七成源于编造或过时信息——幻觉是第一风险源。

重灾区

价格 · 时间 · 开放状态

门票价格、班次时刻、临时闭园——三类信息错得最多，恰恰是行程单的核心。

新增长

版权与肖像纠纷

AI 生成图涉及真人肖像、知名 IP 的争议快速上升——商用素材成重灾区。

隐性区

数据泄露难察觉

上传客人名单、内部报价引发的泄露不易即时发现——发现时往往已造成损失。

两类风险：看得见与看不见

显性风险

翻车能看见

幻觉事实、照骗翻车、文化冒犯——**当场或短期内暴露**，损失直观（投诉、赔偿、口碑）。

隐性风险

泄露无警报

数据外泄、版权埋雷、能力空心化——**长期累积后爆发**，损失更大且难以挽回（法律、竞争力）。

管理策略不同：显性风险靠**核验流程**拦截，隐性风险靠**操作规范**预防——本章后半逐一给出工具。

四大挑战：一张防御地图

挑战①

幻觉

事实风险

编造景点、过时价格、虚构政策——**防御：核验五步法**（PART 02）。

挑战②

隐私

数据风险

客人信息、商业机密误入输入框——**防御：脱敏四法 + 分级**（PART 03）。

挑战③

版权

法律风险

AI 生成物权属模糊、素材侵权——**防御：素材规范 + 标识**（PART 04）。

挑战④

依赖

能力风险

思考外包、判断力退化——**防御：无 AI 时段 + 输出习惯**（模块二 P01 已讲）。

总纲

权责

责任意识

"AI 说的 " 不是免责理由——**发布者对内容负全责**，这是所有防御的地基。

四大挑战对应四套工具——**本讲按 " 识别风险→给出工具→实战演练 " 逐一展开。**

风险分级：三档应对策略

低风险 · 内部草稿

大胆用

几乎零成本

- 自己看的笔记初稿
- 内部头脑风暴
- 非正式练习——错了改就行

中风险 · 对外内容

过流程

核验 + 标识

- 公众号推文、宣传物料
- 事实核验 + 脱敏检查
- AI 辅助标识——三分钟流程

高风险 · 合同承诺

人工兜底

零容忍

- 合同条款、报价单
- 政策引用、安全承诺
- AI 只做草稿，逐字人工审定

分级原则 按“错了损失多大”定级别——读者越多、承诺越硬，把关越严；一把尺子量所有内容，要么过度要么不足。

风险图谱：四问四答

Q1 应用挑战分几类？

技术侧（幻觉 / 黑箱 / 时效）、使用侧（隐私 / 版权 / 依赖）、环境侧（法规完善中）——**八成风险在使用侧。**

Q2 显性与隐性风险区别？

显性**当场暴露**靠流程拦截；隐性**累积爆发**靠规范预防——后者损失更大更难挽回。

Q3 四大挑战是哪些？

幻觉（事实） · 隐私（数据） · 版权（法律） · 依赖（能力）——各配一套防御工具。

Q4 风险分级的标准？

看 "**错了损失多大**"：内部草稿大胆用、对外内容过流程、合同承诺人工兜底。

02

PART 1

幻觉 防御

HALLUCINATION 1

① 三副面孔 ② 高发场景 ③ 核验五步 ④ 三道防线

幻觉的三副面孔

面孔一 无中生有

编造不存在的景点、餐厅、航线——**名词是真的、组合是假的**："云栖坞古村"式虚构。

面孔二 张冠李戴

把 A 景区的政策安到 B 头上、历史年代错配——**元素都真实、关系是错的**，比无中生有更难识别。

面孔三 过期作废

用训练截止前的旧信息答新问题——价格调了、线路停了、政策改了，**答案曾经正确，现在有害**。

识别要领

冷门信息→ 无中生有高发

多元素关联→ 张冠李戴高发

时效信息→ 过期作废高发

高发场景：哪里错得最多

高危三区

价格类（门票 / 房费 / 团费）、**时刻类**（班次 / 开闭园 / 演出）、**状态类**（维修中 / 停业 / 需预约）——行程单的顶梁柱，恰是幻觉最爱。

中危区

历史沿革、民俗解释、地理数据——**大方向对、细节存疑**；写进讲解词前要对照权威资料。

低危区

文案润色、结构建议、创意发散——**不涉及事实**，幻觉无从谈起，放心用。

一句话判断

这条信息**会不会变**？

会变（价格 / 时间）→必核

不易变但**易错**（史实）→抽查

纯观点创意 →放心用

规律：**越“硬”的信息越要核，越“软”的内容越自由**——把核验力气花在刀刃上。

核验五步法：把住事实关

STEP 1

圈出来

标记事实点

把 AI 产出中**所有可证伪的信息**圈出来：名称、价格、时间、数字、政策——先知道要核什么。

STEP 2

分主次

定核验级

影响行程成败的（班次 / 预约）**必核**；一般参考的抽查即可——精力有限，分级处理。

STEP 3

找源头

官方优先

核验顺序：**官方渠道（官网 / 官微 / 电话）→权威平台（地图 / OTA）→实地确认**——转帖不算源头。

STEP 4

交叉验

双源确认

关键信息**至少两个独立来源**对上——单一来源再像真的也要存疑。

STEP 5

留痕迹

核验记录

核过的信息**记下来源与时间**——出问题能快速定位、快速更正。

五步不是形式主义——它把 " 凭感觉信 AI " 变成 "**按流程放行**"，出错率从碰运气变成可控。

快速核验 vs 深度核验

快速版

两分钟

适用：日常内容、时效性强。动作：**地图定位（存在吗） + 官方号翻一眼（营业吗）**——两步覆盖八成翻车点。

深度版

十五分钟

适用：合同引用、媒体报道、付费攻略。动作：**五步全走 + 双源交叉 + 记录留痕**——宁可慢，不可错。

选择标准：**错了损失多大**——发朋友圈用快速版，收粉丝钱的攻略必须深度版。

核验实战：一份行程单的放行

CASE · 五步法走一遍

任务：AI 生成的 " 苏州两日亲子行程 "，发出前核验

圈出事实点 11 处：4 个场馆、3 个班次、2 项价格、2 条政策

核验结果 9 处无误；2 处踩雷：博物馆需预约、游船冬季停运

- **踩雷一**：AI 写 " 现场购票即可 "——官方公众号明示 " 需提前一天预约 "——**预约制信息必须查当期政策**。
- **踩雷二**：游船项目未标季节性——地图评论证实 "11 月至次年 3 月停运 "——**季节性项目要问淡旺季**。
- **修正发布**：两处更正后发出，行程单附 " 核验日期 "——**给自己留了责任边界**，也给读者时效提示。

效率

全程核验 18 分钟——**拦下两处硬伤**，避免了发布后逐条道歉的被动。

方法

两处雷都有共性：**时效类信息**。凡涉及预约、季节、临时政策，一律当期核验。

幻觉高危信息速查表

信息类型	危险度	核验渠道	核验强度
门票价格 / 优待政策	高	官方公众号 / 官网	当期必核
班次与运营时间	高	官方时刻表 + 电话	当期必核
预约 / 限流要求	高	官方预约平台	当期必核
小众点位推荐	高	地图定位 + 实地	双源确认
历史人文知识	中	权威出版物 / 馆方	抽查对照
文案 / 创意 / 结构	低	无需核验	放心使用

建议把这张表**打印贴在工位**——前三类是行程类内容的命门，核验强度最高。

幻觉防御三道防线

01

事前

PREVENT

提示词设防

提示词加约束："不确定的信息请标注；查不到就说查不到"——从源头减少幻觉产出。

02

事中

VERIFY

五步核验

按核验五步过事实点——圈出、分级、找源、交叉、留痕，流程化拦截。

03

事后

CORRECT

更正机制

发现错误**立即更正并注明**——留核验日期、开纠错通道，把损失截断在最小。

防线逻辑 事前减量、事中拦截、事后止损——**三道防线全开，幻觉就从事故降级为可控成本。**

幻觉防御：四问四答

Q1 幻觉的三副面孔？

无中生有（编造）、**张冠李戴**（错配）、**过期作废**（旧信息）——第三种最常见。

Q2 哪里错得最多？

价格、时刻、状态三类 "**会变的信息**"——恰是行程单的顶梁柱，当期必核。

Q3 核验五步法？

圈出来→分主次→找源头→交叉验→留痕迹——把 "**凭感觉信**" 变成 "**按流程放行**"。

Q4 快速版与深度版怎么选？

看错了损失多大：日常内容两分钟快速版，收费内容 / 合同引用十五分钟深度版。

03

PART II

数据安全

DATA SECURITY

① 泄露代价 ② 脱敏四法 ③ 分级策略 ④ 安全守则

数据泄露：看不见的账单

随手一贴的 " 便利 " 0 元

天价

一次客人信息泄露的代价：赔偿、处罚、品牌信任三连击

便利的真正价格

代价①

法律责任 《个人信息保护法》下，泄露客人身份信息可面临**整改、罚款乃至诉讼**—— " 我不知道 " 不是抗辩理由。

代价②

客户流失 游客发现信息被随意处理，**用脚投票且不再回头**——旅游消费高度依赖信任。

代价③

平台处罚 OTA 与内容平台对数据违规商家**降权、下架、清退**——线上生意的根基动摇。

认知纠偏：把资料贴进 AI 对话框 = **把资料交给一家陌生公司的服务器**——这不是操作习惯问题，是数据出境问题。

两类敏感信息：分开对待

客人个人信息

红线信息

身份证号、护照、手机号、住址、行程轨迹——**法律明令保护**：脱敏是底线，不上传公有工具是纪律。

企业商业信息

高危信息

未发布方案、报价体系、客户名单、内部数据——**合同与竞争法保护**：先关 " 数据训练 " 开关，再谈使用。

共同原则：**先问 " 这条信息泄露了会怎样 "**—— 答案严重的，脱敏或干脆不给 AI 。

脱敏四法：给信息“打码”

脱敏工具箱

四个动作

脱敏不是删除信息，是**保留结构与去掉身份**——AI 照样能干活。

四个动作覆盖日常九成场景。

01

代称替换

姓名→“张先生”，公司名→“某地接社”——**身份剥离，任务照跑。**

02

数字打码

手机号 138****、证件前 3 后 4——**保留格式，抹掉实体。**

03

结构抽象

真实行程→“3 晚 4 天、两城联游”——**留框架细节，去可识别信息。**

04

分段投喂

大文档拆块、敏感段人工处理——**敏感的不过机器。**

分级使用：什么数据用哪个工具

公开数据

已发布的宣传物料、公开攻略、行业报告——**公有云工具放心用**，唯一动作：关闭数据训练开关。

内部数据

未发布方案、经营数据——**企业版 / 私有化部署**优先；必须用公有工具时，脱敏后再上。

红线数据

客人身份信息、核心机密——**不上任何 AI 工具**：相关任务由 AI 出框架，真实数据人工填充。

分级口诀

公开的**关开关**就用

内部的**脱敏后**再用

红线的**不过机器**

——AI 出框架，人填真数

分级不是束缚效率——**是让你在出事时能证明 " 已尽合理义务 "**，这是职业安全的护身符。

脱敏实战：一份客诉分析的安全完成

CASE · 脱敏示范

任务：让 AI 分析 20 条客诉找出共性问题 (含客人姓名电话)

危险做法 整表直接贴入对话框——客人信息进入第三方服务器

安全做法 三分钟脱敏后投喂——同样得到分析结论

- **第一步**：删姓名电话列，用 "客 1- 客 20" 编号——**分析不需要知道是谁。**
- **第二步**：房号改 "高级房 01"，日期只留月份——**结构在，身份无。**
- **第三步**：投喂分析 + 人工回填核对——**结论照样准，风险归零。**

效果

脱敏花了 3 分钟，分析质量几乎无损——**安全与效率可以兼得。**

习惯

把脱敏做成**肌肉记忆**：见敏感信息，先想 "这栏 AI 需要吗"——多数时候不需要。

数据安全四条守则

第一条

输入前自问

这条信息泄露了会怎样？——**答案严重的，脱敏或不给。**

第二条

关闭训练开关

新工具第一件事：找到并关闭 "**数据用于模型改进**" 选项。

第三条

分级用工具

公开随意、内部脱敏、**红线不过机器**——按数据敏感度选工具。

第四条

截图先扫一眼

分享 AI 对话截图前，**检查有无客户与内部信息**——截图泄露最高频。

底线

一句话记牢

把资料贴给 AI = 交给陌生服务器；

职业操守从输入框前的三秒犹豫开始。

课堂辩论：安全规范拖慢效率？

有同学说：脱敏、核验、分级太麻烦，"别人都直接贴"。也有同学说：一次事故就够你喝一壶。你怎么看？

—— 议题背景

效率派 · 流程要轻

- 脱敏核验动辄几分钟，**小事大办**
- 行业里 " 直接贴 " 的人多了——**没见谁出事**
- 规范太重，大家干脆**偷偷用、更不透明**

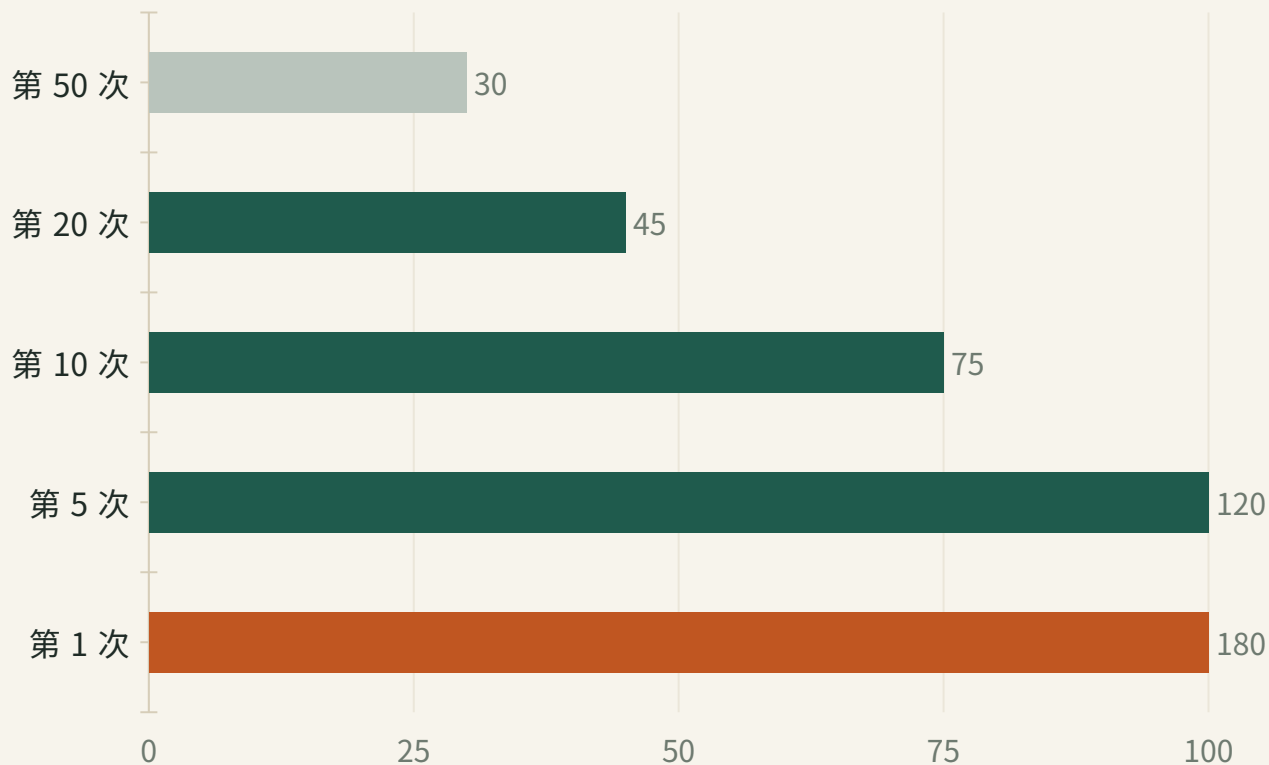
VS

安全派 · 风险要防

- 没出事 ≠ 不会出事——**泄露是概率事件**，累积必然
- 法规趋严，**秋后算账**的案例已出现
- 流程熟练后只需几十秒——**成本被高估了**

讨论规则：用 "**错了损失多大 × 发生概率**" 给三条日常操作定价。**辩完自问：你的定价理性吗？**

安全习惯养成曲线（示意）



示意数据（单位：秒），综合课堂演练记录绘制，仅作参考。

01

第一次脱敏最慢（约 3 分钟）——**陌生动作的天然成本。**

02

十次之后进入**半分钟区间**——流程内化为习惯。

03

熟练后成本趋近于零——**安全**的“慢”是暂时的，事故的“贵”是**永久的。**

04

PART F

合规 生产

COMPLIANT CR

① 合规三层 ② 内容标识 ③ 素材规范 ④ 发布清单

合规三层：从内容到流程

L1 内容层 真实无害

事实核验通过、无虚假宣传、无文化冒犯——**内容本身的合规**，幻觉防御的主战场。

L2 标识层 AI 使用透明

按规标识 AI 生成 / AI 辅助——**告知义务的合规**：读者有权知道内容的人机比例。

L3 流程层 可追溯

核验记录、脱敏流程、素材来源留痕——**管理动作的合规**：出事能自证尽责。

三层关系 内容层防事故、标识层防争议、流程层防追责——**三层齐备，才算真正的合规生产。**

内容标识： AI 使用要亮牌

为什么必须标识

《生成式人工智能服务管理暂行办法》要求对 AI 生成内容**显著标识**；各内容平台同步上线 AI 标识规则——**合规义务，不是可选动作**。

怎么标识

平台自动标识之外，自制内容建议文末注明：**" 本文由 AI 辅助生成，经人工核验 "**——一句话，信任与合规兼得。

标识的意义

不是自贬身价——**是告诉读者 " 我用了工具，但我负全责 "**；先行标识者反而在信息浑水中赢得信任。

标识话术

" 本文 AI 辅助生成，信息经人工核验 "

" 图片由 AI 生成，仅作参考 "

" 数据来源：XX 统计年鉴（AI 整理） "

拿不准标不标时——**标**。多标的代价几乎为零，漏标的代价可能很贵。

素材使用规范速查表

素材类型	风险点	使用规范
AI 生成配图	权属模糊 · 细节失真	商用注明 AI 生成；核验文字与细节
真人肖像素材	肖像权纠纷	用授权图库；AI 生成脸避开真人像
知名 IP 元素	商标 · 著作权侵权	不用未授权 IP ；避开 "XX 风格 " 提示词
实拍风景图	拍摄者著作权	自有拍摄或授权图库；注明来源
音乐音效	平台版权检测	用平台曲库或免版权音源

总原则：**能溯源的素材才是安全素材**——来源不明的 " 免费图 " ，往往最贵。

事故复盘：一张 AI 海报引发的索赔

CASE · 教训

事件：某景区用 AI 生成的宣传海报，人物神似某明星（综合公开报道改编）

经过 海报投放三天后，被指“蹭明星肖像”，当事人发函

结果 全线下撤 + 公开致歉 + 协商补偿——半个月的宣发白做

- **根因一**：提示词写“明星感美女”——AI 直接向真实人脸靠拢，**提示词本身就是风险源**。
- **根因二**：发布前无人审视相似度——**审核流程缺位**，没有“像谁”这一问。
- **正确姿势**：人物海报用**明确虚构设定**（“原创虚拟人物，非真实人像”）+ 发布前相似度自查。

教训

AI 生成“像谁”的责任**在使用者**——“是 AI 自己画的”不是抗辩理由。

改进

涉人素材增设一道“**相似度三问**”：像谁？像到什么程度？换成谁？

标识三场景：亮牌的学问

场景① · 纯 AI 生成

明示生成

全亮牌

- " 本图由 AI 生成 "
- 配图示意类标注用途
- 无事实声明风险时适用

场景② · AI 辅助创作

注明协作

半亮牌

- "AI 辅助生成，人工核验 "
- 列出人机分工边界
- 多数工作内容适用

场景③ · 数据整理

溯源标注

亮来源

- " 数据来源：官方统计 "
- AI 只做整理不产出观点
- 报告类内容必备

共同底线 三种亮法，一个内核：读者有权知道内容的 " 人机配方 "—— 透明是最省成本的合规。

发布前检查清单：六问放行

| 内容对外发布前，过一遍这六问——三分钟，拦下九成风险：

01 事实核了吗 价格 / 时间 / 预约类信息当期双源核验

02 数据脱敏了吗 客人信息、内部数据零残留

03 素材可溯源吗 图片音乐有授权来源，肖像已自查

04 标识加了吗 AI 生成 / 辅助按场景亮牌

05 文化审了吗 民族宗教元素、地域梗过 sensitivity 一遍

06 留痕了吗 核验记录、素材来源存档备查

四个问题，四个答案

Q1 四大挑战与防御？

幻觉→核验五步；隐私→脱敏分级；版权→素材规范 + 标识；依赖→输出习惯——**风险地图对应工具箱**。

Q2 核验五步法？

圈出来→分主次→找源头→交叉验→留痕迹——把 " 凭感觉信 " 变成 " 按流程放行 "。

Q3 数据安全的核心动作？

输入前自问泄露后果 + **脱敏四法** + 分级用工具——红线数据不过机器。

Q4 发布前最后一关？

六问清单：事实、脱敏、溯源、标识、文化、留痕——全过再发。

作业：一次全流程合规实践

思考题 · 下次课开篇讨论

- ① 对照四大挑战，你过往使用中最欠账的是哪一项？打算怎么补？
- ② "六问清单"里，哪一问在你看来最容易被跳过？为什么？

课后作业 · 实践任务

任务一

选一份 AI 生成的旅游内容（自己产出或网络案例），完成核验五步 + 发布前六问，提交核验记录表。

重点记录：发现了几处问题、各属哪类风险、如何修正。

任务二

为你常接触的一类业务（行程单 / 推文 / 海报）定制一份 "脱敏 + 核验" 迷你流程卡。

A5 卡片大小，三步以内——好用比全面重要。

提交：**下次课前**上传课程平台；优秀流程卡将印制为班级共用工具。

延伸学习资源

法规

《生成式人工智能服务管理暂行办法》全文

二十四条不长——**通读一遍**，标识义务与红线心中自明。

指南

各内容平台 AI 标识规则说明

你发布的每个平台都有一份——**收藏夹建一个 " 合规 " 文件夹**。

案例

AI 内容纠纷公开判例与通报

看真实代价——**案例的震撼胜过十条规范**。

工具

图片版权检测与溯源工具

发图前扫一遍——**让工具帮你看住盲区**。

模块二收官。下一模块进入**八大场景应用**：行程规划、导游词、新媒体、营销、翻译、客服、数字人、数据分析。

模块三 · 项目一：个性化行程规划

MODULE 03 · PROJECT 01

场景， 开战

从 " 会工具 " 到 " 出方案 "

01 需求采集

把模糊的 " 想去玩 " 翻译成**结构化需求**。

02 AI 规划流

提示词模板 + 迭代打磨，**半日出初稿**。

03 人工精修

核验、取舍、加入**人的温度**——成品的最后一公里。