2025 / 12

中国AI应用系统
合规参考手册

帮助 AI 产品项目团队快速理解合规框架与关键动作
有效规避项目落地的合规风险

免责声明：本手册基于公开资料与实践经验整理，仅供参考，不构成法律建议。具体合规事项建议咨询专业法律顾问。

合规框架总览

中国AI治理体系：法律 → 监管办法 → 技术标准 → 认证备案

1 法律层

原则与责任

强制

网络安全法
数据安全法
个人信息保护法

2 监管办法

AI场景细化监管

强制

生成式AI管理办法
算法推荐管理规定
深度合成管理规定

3 技术标准

工程实现规范

推荐/强制

GB/T 35273（个人信息安全）
GB/T 22239（等保基本要求）
GB/T 45654（生成式AI安全）

4 认证备案

落地合规手段

强制

ICP/App备案
等保测评
AI大模型/算法备案

★ **开发AI应用前，建议先对照此框架判断自身所处阶段与适用监管层级**

主要法律法规速查

法规名称	适用对象	核心要求	强制/推荐	实施时间
网络安全法	所有互联网平台	等保制度、安全义务、关键信息基础设施保护	**强制**	2017
数据安全法	数据处理者	数据分类分级、安全管理制度、风险评估	**强制**	2021
个人信息保护法	处理个人信息的企业	用户同意、敏感信息保护、跨境传输限制	**强制**	2021
生成式AI管理办法	AI大模型/内容平台	模型备案、数据合法来源、内容安全	**强制**	2023
算法推荐管理规定	推荐系统平台	算法备案、禁止算法歧视	**强制**	2022
深度合成管理规定	AI换脸/语音/数字人	深度合成备案、内容标识	**强制**	2023
AI生成内容标识办法	所有生成式AI平台	AI生成内容必须标识来源	**强制**	2025
网络数据安全管理条例	互联网企业	数据处理规范、数据出境、安全审计	**强制**	2025

★ 以上均为强制执行法规，AI产品上线前须逐项对照评估

核心法律详解（一）

网络安全法 · 数据安全法 · 个人信息保护法

网络安全法

2017

适用：所有互联网平台 / 信息系统运营者

- 等级保护制度：信息系统须按重要程度定级并通过测评
- 安全保护义务：需建立安全管理制度、技术防护、应急预案
- 关键信息基础设施：重要行业需额外保护，数据须境内存储
- 网络安全事件：发生重大事件须立即向主管部门报告

⚠ 未落实等保可面临停业整顿，情节严重最高100万元罚款

数据安全法

2021

适用：数据处理者（含AI训练数据使用方）

- 数据分类分级：按重要程度对数据进行分类并实施差异化保护
- 数据安全审查：影响国家安全的数据处理活动须通过安全审查
- 数据交易规范：数据提供方须审核数据来源合法性
- 数据出境限制：重要数据出境须经安全评估

⚠ 违规处理重要数据最高可处200万元罚款；情节严重可吊销营业执照

个人信息保护法

2021

适用：所有处理个人信息的企业（AI产品几乎均适用）

- 知情同意原则：处理个人信息须告知目的并取得明确同意
- 最小必要原则：收集信息应限于实现目的的最小范围
- 敏感信息保护：生物识别、健康、金融等须单独授权
- 用户权利保障：须支持用户查阅、复制、更正、删除个人信息
- 个人信息影响评估（PIA）：自动化决策场景须事先评估

⚠ 违规最高罚款5000万元或年营业额5%；情节严重暂停/终止服务

核心监管办法详解（二）

生成式AI管理办法 · 算法推荐管理规定 · 深度合成管理规定

生成式AI服务管理暂行办法

2023

适用：面向中国公众提供生成式AI服务的企业

- 服务备案：正式上线前须完成大模型备案
- 数据合法性：训练数据须有合法来源，不得侵犯知识产权
- 内容安全：须建立内容审核机制，防止违法有害信息生成
- AI标识：AI生成内容须有显著标识
- 用户协议：须制定明确的用户服务协议和隐私政策
- 安全评估：上线前须进行安全评估

⚠ 未备案擅自提供服务，可责令停止服务、没收违法所得

互联网信息服务算法推荐管理规定

2022

适用：使用推荐算法向用户提供个性化内容的互联网平台

- 算法备案：具有舆论属性或社会动员能力的推荐算法须备案
- 禁止算法歧视：不得根据用户特征设置差异化价格
- 关闭权利：须向用户提供关闭个性化推荐的选项
- 信息透明：须向用户说明推荐逻辑
- 未成年人保护：须对未成年人提供专门保护

⚠ 违规可处5万~50万元罚款；情节严重责令暂停相关业务

互联网信息服务深度合成管理规定

2023

适用：提供AI换脸、语音克隆、虚拟数字人等深度合成服务的企业

- 服务备案：深度合成服务提供者须在网信办进行备案
- 内容标识：深度合成内容须添加显著标识，不得误导公众
- 真实身份核验：须对使用者进行实名认证
- 禁止用途：不得利用深度合成制造虚假新闻、欺诈内容
- 数据安全：人脸/声纹等生物特征数据须依法处理

⚠ 未备案或未标识，可处5万~50万罚款；情节严重暂停/终止服务

技术标准速查

以下标准为工程落地参考，推荐标准不具强制约束力，但通常作为监管评估依据

GB/T 35273 2018

个人信息安全规范 个人信息保护最佳实践、隐私政策、用户同意

GB/T 22239 2019

等保基本要求 网络安全等级保护安全技术基线要求

GB/T 28448 2019

等保测评要求 信息系统安全等级保护测评方法

GB/T 37964 2020

个人信息去标识化 数据脱敏和匿名化技术规范

GB/T 45654 2025

生成式AI安全基本要求 生成式AI服务安全基线（2025新标准）

GB/T 45288 2024

大模型技术要求 大模型能力定义与技术指标

GB/T 45674 2025

AI数据标注安全规范 训练数据标注安全流程

AI内容标识规范 2025

内容水印与元数据标识 AI生成内容水印技术规范

AI产品合规决策树

根据产品类型快速定位所需合规动作

是否对公众提供互联网服务?

否 — 企业内部AI

等保定级与测评
（等保2级）

个人信息保护合规
隐私政策 / PIA

数据安全治理
分类分级 / 访问控制

数据出境合规
如涉及跨境数据

个人信息相关（通用）

- 隐私政策（完整、清晰）
- 用户授权机制
- 数据最小化原则
- 个人信息影响评估（PIA）
- 用户权利保障（访问/删除/纠错）

是 — 互联网AI产品

ICP / 公安 / App 备案
上线前必须完成

等保测评
通常等保2级或3级

生成式AI备案
如生成AI内容

算法/深度合成备案
如推荐算法/换脸/数字人

合规动作全览（5个阶段·21项关键动作）

① 产品规划	② 研发阶段	③ 上线准备	④ 上线阶段	⑤ 运营阶段
AI产品合规评估	数据治理体系建设	ICP备案 必须	生成式AI备案 必须	内容审核机制
数据来源合法性评估	数据脱敏与匿名化	公安备案 必须	算法备案 必须	用户投诉与纠错
数据分类分级	AI内容安全机制	App备案 必须	深度合成备案 必须	算法透明机制
个人信息影响评估（PIA）	AI生成内容标识	等保定级与测评	数据出境安全评估	数据安全监控
	隐私保护设计	AI备案材料准备		数据安全事件响应

红色边框 = 必须完成（上线硬性要求）

※ ISO 27001 等认证详见「认证体系专页」

认证体系专页

强制测评 + 推荐认证 — AI产品合规能力建设的重要支撑

强制性测评（上线必须）

等保 2/3 级

等级保护测评

按网络安全等级保护制度，信息系统须通过等保测评方可正式运营。AI产品一般为2级，处理大量个人信息或重要业务为3级。

周期：1.5 ~ 3 个月

主管：公安部认定的测评机构

国家评估

数据出境安全评估

向境外提供重要数据，或处理100万人以上个人信息、跨境提供10万人个人信息等情形须通过国家安全评估。

周期：3 ~ 6 个月

主管：国家互联网信息办公室

推荐认证（能力建设 / 增信）

国际认证

ISO/IEC 27001

信息安全管理体系

全球通行的信息安全管理体系认证。建立系统化的安全管理制度，覆盖风险评估、访问控制、事件响应等领域。

价值：合作伙伴/客户认可度高；可作为投标/上市合规依据；减少安全审查频率

周期：6 ~ 12 个月

国际认证（2023新）

ISO/IEC 42001

AI管理体系

专门针对AI系统的管理体系国际标准。涵盖AI风险管理、透明度、问责制等要求，与国内生成式AI监管要求高度契合。

价值：向监管机构展示AI治理能力；出海业务的重要信任背书

周期：6 ~ 12 个月

国内认证

个人信息保护认证

GB/T 35273

由国家认证认可监督管理委员会主管，按GB/T 35273标准对企业个人信息保护能力进行认证评估。

价值：可替代部分数据出境标准合同场景；国内用户信任建设

周期：3 ~ 6 个月

建议路径：上线前完成「等保测评」（强制）→ 上线后12个月内启动「ISO 27001」（强烈推荐）→ 视业务场景补充「ISO 42001 / 个人信息保护认证」

关键备案事项详解

以下备案均为强制性要求，缺一不可

备案类型	主管部门	备案时机	适用对象	办理方式简述	风险提示
ICP 备案	工信部	**上线前**	提供网站/Web服务的企业	通过接入商（阿里云/腾讯云等）提交，通常需2-4周	未备案不得开展互联网信息服务，面临关站处罚
公安备案	公安机关	**上线后30天内**	网站/Web平台	全国互联网安全管理服务平台（beian.mps.gov.cn）在线提交	未备案面临罚款及停止服务处罚
App 备案	工信部	**上线前**	移动应用（iOS/Android）	通过应用分发平台或工信部App备案系统提交	2024年起强制执行，未备案App将被下架
算法备案	网信办	**上线前**	具备舆论属性或社会动员能力的推荐算法产品	网信办算法备案系统在线备案	未备案不得提供推荐服务
生成式AI备案	网信办	**上线前**	对中国公众提供生成式AI服务的大模型/AI内容平台	网信办生成式AI服务备案，需提供模型说明、安全评估报告	未备案不得向公众提供生成式AI服务
深度合成备案	网信办	**上线前**	AI换脸、语音克隆、虚拟人等深度合成服务提供者	网信办深度合成服务备案	未备案不得提供深度合成服务

等保测评 — 关键合规里程碑

等保（网络安全等级保护）是中国信息系统安全的基础制度，将系统按重要程度分为1-5级，AI产品通常需达到 2级 或 3级 要求。

1级 用户自主保护

适用： 一般个人用户系统

AI产品： 通常不适用

2级 指导保护

适用： 一般互联网产品

AI产品： 多数AI应用

3级 监督保护

适用： 涉及大量个人信息或重要业务

AI产品： 金融/医疗/大规模AI

4级 强制保护

适用： 关键信息基础设施

AI产品： 极少数AI产品

测评流程（通常耗时 1.5 - 3 个月）

1. 系统定级（自评）
2. 定级备案（公安机关）
3. 差距分析（整改评估）
4. 安全整改（技术/管理）
5. 等保测评（测评机构）
6. 出具报告（通过/不通过）

⚠ **建议上线前 3-4 个月启动等保定级，预留足够整改与测评时间**

合规时间轴 — 项目各阶段关键节点

D-180天+ 项目启动

- AI合规评估
- 数据来源审查
- 数据分类分级

D-120天 研发启动

- 数据脱敏设计
- AI安全机制设计
- 隐私保护设计

D-90天 上线准备

- 等保定级启动
- ICP备案申请
- 公安备案申请

D-30天 内测/灰度

- AI备案材料准备
- 内容审核机制验证
- 等保测评通过

D-Day 正式上线

- 生成式AI备案✓
- 算法/深度合成备案✓
- 所有备案完成✓

⚠ 备案与等保测评周期较长（1-3个月），务必在项目计划中前置预留，切勿在上线前临时启动

上线后 — 运营阶段持续合规

合规不是一次性任务，而是持续运营的基本能力

审 内容审核机制

高

AI输出内容须建立敏感词过滤、违规内容识别、人工复审等多层机制，覆盖文本、图像、语音全模态

诉 用户投诉与纠错

高

必须提供用户申诉渠道，对AI错误信息建立纠错机制，且须在合理期限内响应，不得拒绝或拖延

透 算法透明机制

高

推荐算法平台须向用户提供关闭个性化推荐的选项，并以用户可理解的方式说明推荐逻辑

监 数据安全监控

高

持续监控数据访问行为、异常操作，建立安全事件预警机制，定期进行数据安全审计

应 数据安全事件响应

高

建立数据泄露应急预案，发生重大安全事件须在规定时间内（通常72小时）向主管部门报告

建议每半年开展一次合规自查，关注监管政策更新，及时调整合规措施

合规行动清单 — 项目团队自查

产品规划

- □ 判断产品类型（生成式AI / 推荐算法 / 深度合成）
- □ 完成数据来源合法性评估
- □ 建立数据资产分类分级清单

研发阶段

- □ 设计AI内容安全机制（敏感词过滤等）
- □ 完成训练数据脱敏与匿名化
- □ 实现AI生成内容标识功能
- □ 完成隐私政策与用户授权机制设计

上线准备

- □ ICP备案申请已提交
- □ 公安备案已提交
- □ App备案已完成（如适用）
- □ 等保定级已完成，测评已启动

上线阶段

- □ 生成式AI备案已通过
- □ 算法备案已通过（如适用）
- □ 深度合成备案已通过（如适用）

运营阶段

- □ 内容审核机制已上线
- □ 用户申诉/纠错渠道已开放
- □ 数据安全监控已部署

认证（推荐）

- □ 等保测评已通过（强制）
- □ ISO 27001 认证（强烈推荐）
- □ ISO 42001 / 个人信息保护认证（视业务）