生成式人工智能服务合规备案指南

（2026 年）

2026 年 1 月

中国·北京

为帮助生成式人工智能服务提供者系统识别并有效应对合规备案过程中的法律风险，探索契合生成式人工智能技术特征与监管要求的可操作路径，在深入调研现行监管政策、备案实践与企业合规需求的基础上，编制形成《生成式人工智能服务合规备案指南》。

《生成式人工智能服务合规备案指南》共分为五个部分，系统构建生成式人工智能服务备案的制度认知、操作路径与风险防控框架。

围绕生成式人工智能服务备案的制度背景、操作路径与风险防范，本指南从制度认知—流程操作—风险控制—工具支持四个层面展开，力求为企业提供一套可理解、可操作、可复用的合规备案指引体系。

第一部分从监管制度出发，明确指南的编制目的、基本原则及核心术语，界定生成式人工智能服务备案的适用范围与合规边界，为企业准确理解备案要求提供基础认知。

第二部分围绕生成式人工智能服务备案的发展现状，系统梳理中央及地方层面的政策文件与制度安排，并结合实践情况分析当前备案制度的运行特点与监管重点，帮助企业判断自身服务在现实监管中的定位。

第三部分聚焦生成式人工智能服务备案的实务操作，详细解析备案的具体要求、流程安排及常见难点，厘清备案、登记等不同合规路径之间的关系，引导企业提前开展针对性合规准备。

第四部分从风险防控视角出发，系统梳理未依法履行备案义务的主要类型及其可能承担的法律责任，并总结企业在备案实践中的常见错误与经验教训，提示生成式人工智能服务企业需要重点防范的合规风险。

第五部分作为工具性支持内容，汇集生成式人工智能服务备案相

关的重要法律法规、政策文件、技术标准及流程示例，为企业具体开展备案与合规工作提供参考依据。

本指南得到北京市文化娱乐法学会、北京中周法律应用研究院、北京法科智情信息技术有限公司的支持。指南由智情日参团队撰写，成员包括崔星璐、范艺颖、魏日升、刘磊、汪慧玲、郑祖星、杨鸿、王文斌、徐少明、邓辉、潘嘉斌、席崇俊、王珊珊、曲占峰、李凡、吕晓擎、王天宇、严超亮、郭恩泽、徐峥玮、陈业昕、文睿鸣、史佳潞、王梓淼、涂思羽。

我们希望通过本指南的发布，为我国生成式人工智能服务提供者在合规备案与规范运营过程中提供可参考、可落地的公共指引，降低企业合规理解与实践成本，促进技术创新与依法治理的良性互动。作为一份面向社会开放的公益性研究成果，指南内容旨在服务产业发展与社会公共利益，推动生成式人工智能在开放可控、安全可信的轨道上实现高质量发展。诚邀业界同仁、研究机构与相关主体浏览、参考并提出宝贵意见，共同完善我国人工智能治理实践。

生成式人工智能服务合规备案指南

前 言

当前新一轮科技革命和产业变革深入发展，**人工智能技术正以其显著的创新性、交叉性与前沿性，成为重塑全球经济结构、改变社会治理范式的关键力量。**特别是以大模型为代表的生成式人工智能（Generative AI），通过对算力、算法、数据三要素的深度融合与重构，已从单一的感知智能迈向认知智能的新阶段，成为发展新质生产力的核心引擎。在全球范围内，主要科技强国均将人工智能视为大国博弈的战略制高点，不仅在基础大模型研发、高端芯片制造等技术维度展开激烈角逐，更在法律治理与规则制定层面加速布局，力图掌握未来发展的主动权。我国在“十五五”规划建议的顶层设计中，**明确提出要加快科技自立自强，坚持创新在我国现代化建设全局中的核心地位，统筹发展和安全，**这为人工智能产业在法治轨道上的高质量发展指明了方向，也对构建安全可信的产业生态提出了更高要求。

从技术演进与风险挑战的态势审视，人工智能正处于从技术爆发向全面应用转化的关键窗口期。一方面，模型参数规模的指数级跃升与多模态能力的深度拓展，极大地释放了社会创新活力，赋能千行百业；另一方面，技术应用的“双刃剑”效应日益凸显，带来了前所未有的复杂挑战。随着生成内容的逼真度与传播速度大幅提升，**数据安全泄露、算法偏见歧视、虚假信息生成、知识产权侵权以及伦理道德风险等问题交织叠加，且具有隐蔽性强、破坏力大、传播范围广等特点。**全球人工智能治理格局正经历深刻变革，从单纯的技术竞争转向治理能力竞赛。如何在保障国家主权、安全、发展利益的前提下，构

建全要素、分主体、全周期的风险防控体系，最大限度地激发市场主体活力，实现技术红利与社会福祉的动态平衡，已成为行业发展的必答题。

从我国的政策法规布局审视，国家正加速构建具有中国特色、并在国际上具有重要影响力的人工智能法律治理体系。**贯彻落实《生成式人工智能服务管理暂行办法》等一系列法规政策，不仅是全面推进依法治国的具体体现，更是确立行业规则、稳定市场预期的必然选择。**我国坚持“包容审慎、分类分级”的监管原则，正在逐步完善从算法推荐、深度合成到生成式人工智能的“分层级、分领域”人工智能治理框架。**在这一体系中，生成式人工智能模型备案不是简单的行政管理手段，而是连接技术研发与市场准入的关键枢纽，是落实“源头治理”理念、防范系统性风险的重要抓手，也是保障产业生态健康、有序发展的必要手段。**通过备案机制，监管部门能够及时掌握技术发展动态，企业能够确立合规边界，用户能够获得安全保障，从而形成政府、企业、社会多方协同共治的良好局面。

本指南旨在深入贯彻落实国家关于人工智能发展的战略部署，明确人工智能模型备案的具体要求与操作规范，为人工智能服务提供者、技术支持者及相关从业主体提供清晰、可执行的合规指引。指南坚持问题导向与目标导向相结合，通过**厘清备案范围、规范备案流程、统一材料标准、界定关键术语**，致力于解决当前备案实践中存在的标准不一、流程不明等痛点问题，**切实降低企业合规成本，提升备案效率，**助力生态伙伴在合规的轨道上加速奔跑。

展望未来，坚持合规经营是人工智能企业确立市场主体地位、实现高质量发展的根本前提，也是应对全球化竞争挑战的必由路径。我

们愿与监管部门、行业协会及广大从业者携手并进，共同维护、迈向一个**安全可信、包容审慎、创新活跃**的人工智能产业生态。

目 录

表 目 录

图 目 录

第一章 总则

一、指南目的及基本原则

（一）指南目的

在贯彻落实总体国家安全观，适应全球人工智能治理格局变革，响应国家“十五五”规划关于加快科技自立自强、统筹发展和安全的战略部署的背景下，本指南应运而生。本指南依据现行法律法规，旨在为生成式人工智能服务提供者开展备案工作提供专业指引，解决行业实践中面临的监管标准理解不一、合规操作路径模糊等问题，具体目的如下：

一是解析法律法规要求，细化合规审查标准。依据《中华人民共和国网络安全法》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》《生成式人工智能服务管理暂行办法》等法律法规，梳理生成式人工智能服务提供者在备案环节应当履行的法定配合义务。指南将抽象的法律条款转化为具体的备案材料清单与审查要点，指引从业主体对训练数据来源合法性、算法机制透明度、生成内容安全性进行自我评估与申报，协助企业准确理解并满足监管要求。

二是梳理风险防控体系，辅助构建安全屏障。针对生成式人工智能具有的舆论属性、社会动员能力及内容生成不可控特征，解析监管部门关于事前预防与事中监管的治理逻辑。指南通过拆解数据泄露、算法偏见、虚假信息生成及意识形态安全等风险点的具体的表现形式，指导企业建立健全覆盖模型全生命周期的网络安全与信息安全保护制度，提升企业防范单点技术风险向系统性风险演化的能力。

三是统一备案操作规范，降低合规试错成本。针对当前备案实践中存在的申报材料颗粒度不一、技术评估维度不明确等痛点，提供标

准化、规范化的操作流程指引。指南明确备案适用范围、细化技术支撑材料编写规范、规范安全评估申报流程，旨在消除政策理解偏差，帮助从业主体减少因材料不合规导致的反复修改与退回，提升备案通过效率，促进算力、算法、数据等创新要素在合规前提下顺畅流动。

四是厘清主体责任边界，提供精准备案路径。针对人工智能产业链条长、参与主体多的特点，依据现有法律法规关于“提供者”与“支持者”的定义，梳理不同技术角色在数据处理、内容审核、安全防护环节的责任分工。指南旨在帮助基础模型开发者、应用服务提供者准确识别自身法律定位，选择适配的备案路径与申报策略，避免因责任认知不清导致备案主体错误或责任界定混淆。

（二）指南原则

本指南在编制及指导备案实践过程中，遵循以下基本原则：

一是坚持统筹兼顾，发展与安全并重。贯彻“十五五”规划关于统筹发展和安全的要求，深刻认识发展是安全的基础，安全是发展的保障。备案指引工作依据国家安全底线要求，提示企业对利用生成式人工智能从事违法违规活动的法律风险；同时充分尊重技术发展规律，在解读监管要求时注重平衡性，指导企业在确保安全底线的前提下，充分保障技术创新与发展的空间。

二是坚持源头治理，全要素全周期管控。依据源头治理与全过程监管相结合的治理范式，指引企业将合规动作前移。备案工作重点关注语料库来源的合法性审查、标注规则的科学性设定以及算法机制的安全性评估，协助企业从源头上识别并阻断风险。指导企业建立覆盖模型训练、优化、部署、服务及退出等全生命周期的风险防控证据链，确保在备案审查中能够提供风险可识别、可监测、可处置的有效证明。

三是坚持依法依规，严守合规底线。严格遵循《中华人民共和国网络安全法》《中华人民共和国数据安全法》《个人信息保护法》《生成式人工智能服务管理暂行办法》等现行法律法规，将合法合规作为备案工作的核心准则。指南依据法定程序与标准，指引备案主体全面履行法定义务，建立健全数据安全与内容审核机制。提示企业在备案申报过程中必须保证材料的真实性、完整性与准确性，严禁弄虚作假或隐瞒关键技术细节，确保人工智能技术研发与应用始终在法治轨道上运行。

二、相关术语界定

1.生成式人工智能服务，是指利用生成式人工智能技术向中华人民共和国境内公众提供生成文本、图片、音频、视频等内容的服务。

2.生成式人工智能技术，是指具有文本、图片、音频、视频等内容生成能力的模型及相关技术。

3.生成式人工智能服务提供者，是指利用生成式人工智能技术提供生成式人工智能服务（包括通过提供可编程接口等方式提供生成式人工智能服务）的组织、个人。

4.生成式人工智能服务使用者，是指使用生成式人工智能服务生成内容的组织、个人。

5.算法备案，是指针对提供具有舆论属性或者社会动员能力的算法推荐服务(包括生成合成类、个性化推送类、排序精选类、检索过滤类、调度决策类等算法技术)，需依据《互联网信息服务算法推荐管理规定》等法律法规，应通过中央网信办线上提交《落实算法安全主体责任情况》《算法安全自评估报告》《拟公示内容》等材料。

6.生成式人工智能服务备案，是指具备舆论属性或者社会动员能

力的生成式人工智能服务，需根据《生成式人工智能服务管理暂行办法》等法律法规，应通过省级网信办线下提交《生成式人工智能(大语言模型)上线备案申请表》《安全自评估报告》《模型服务协议》《语料标注规则》《关键词拦截列表》《评估测试题集》等材料。

7.生成式人工智能服务登记，是指对通过 API 接口或其他方式直接调用已备案生成式人工智能服务能力，且面向境内公众提供具有舆论属性或者社会动员能力的生成式人工智能服务，应通过属地网信办提交服务提供者的情况、调用的生成式人工智能服务的情况说明等。

第二章 生成式人工智能服务备案的发展现状

一、中央层面生成式人工智能服务备案的相关政策文件

通过在中央网信办官方网站，以“人工智能”和“备案”为关键词，并对检索结果进行人工筛选，共得到与人工智能备案或算法备案相关的4份法律法规文件。此外，在北大法宝以“人工智能”和“备案”为关键词进行检索，还可检索出国务院颁布的1份法律法规文件，具体内容如下表。

表 1 中央层面人工智能备案相关政策文件

文件名称	施行时间	发布机构	文件获取链接	具体提及处
互联网信息服务算法推荐管理规定	2022.3.1	国家网信办	https://www.cac.gov.cn/2022-01/04/c_1642894606364259.htm	第24条
互联网信息服务深度合成管理规定	2023.1.10	国家网信办	https://www.cac.gov.cn/2022-12/11/c_1672221949354811.htm	第19条
生成式人工智能服务管理暂行办法	2023.8.15	国家网信办	https://www.cac.gov.cn/2023-07/13/c_1690898327029107.htm	第17条
人工智能生成合成内容标识办法	2025.3.7	国家网信办	https://www.cac.gov.cn/2025-03/14/c_1743654684782215.htm	第12条
国务院关于深入实施“人工智能+”行动的意见	2025.8.26	国务院	https://www.gov.cn/zhengce/zhengceku/202508/content_7037862.htm	（三）13

二、地方层面生成式人工智能服务备案的相关政策文件

通过在各省市、区网信办官方网站，以“人工智能”“人工智能服务备案”“若干措施”“补贴政策”为关键词，检索得到以下结果。

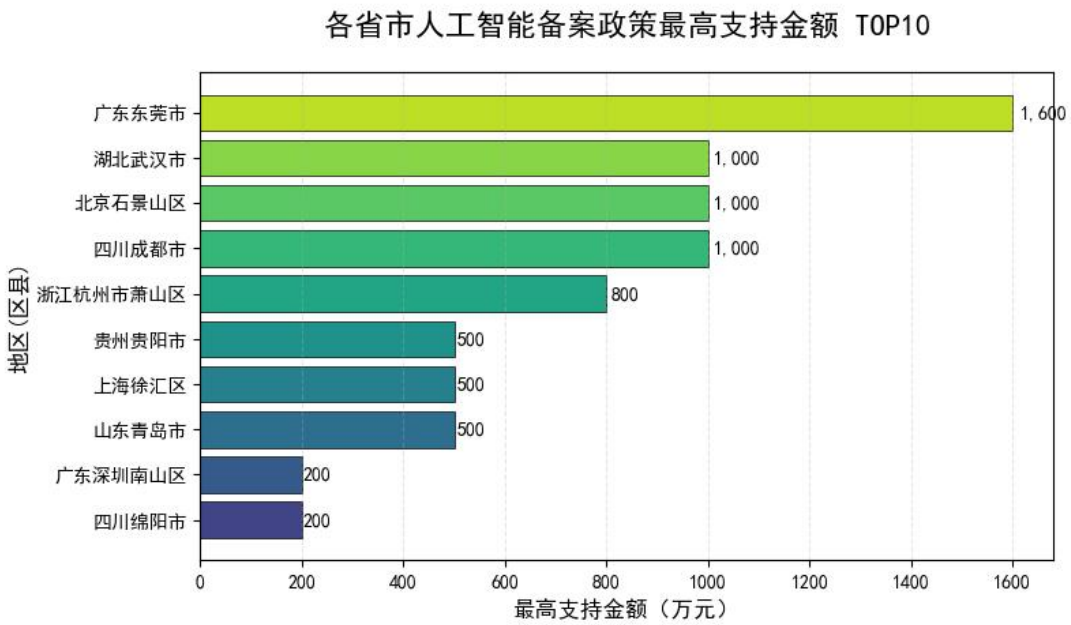


图 1 各省市人工智能备案政策最高支持金额 TOP10

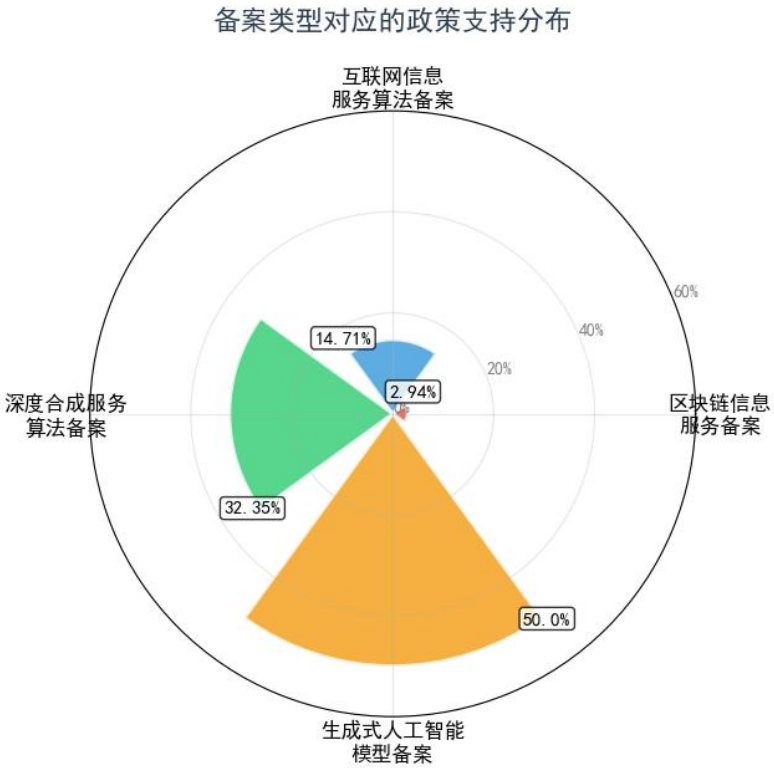


图 2 各备案类型对应政策支持分布

（一）北京市

石景山区：对于通过国家《生成式人工智能服务管理暂行办法》备案的大模型，购买或租用算力的，按照不超过当年实际发生金额30%比例，给予大模型研发主体最高1000万元资金支持。

经济技术开发区：对新增获得国家互联网信息办公室生成式人工智能模型备案的企业，给予一次性100万元支援。

丰台区：（1）支持企业自主研发人工智能大模型，对通过国家算法、深度合成等备案的大模型，一次性给予研发企业最高50万元支持；对于工业领域大模型，一次性给予最高60万元支持。（2）对购买通过国家备案的通用大模型开展自有垂类大模型研发的企业，按照不超过年度购买服务实际投入的30%给予最高20万元支持，连续支持不超过2年。

通州区：支持企业开展软件研发、算法创新、模型攻关、方案创新等，促进产业迭代升级，对新增获得国家互联网信息办公室生成式人工智能服务、境内深度合成服务算法、区块链信息服务备案且落地应用效果良好的企业，分别给予最高100万元、20万元、10万元一次性支持，单个企业每年最高给予100万元支持。

怀柔区：支持人工智能示范应用。围绕智能算力、大模型和应用赋能等人工智能全产业链领域，对于落地的高成长型企业以及获得中央网信办生成式人工智能模型备案的企业，给予一次性50万元资金支持。

表 2 北京市生成式人工智能备案相关政策文件

文件名称	施行时间	发布机构	文件获取链接	具体提及处
石景山区推进	2024.3.19	中关村科技园区	https://www.bjsjs.g	第四条

人工智能大模型产业集聚区建设发展支持办法		石景山园管理委员会北京市石景山区科学技术委员会	ov.cn/gongkai/zwgkpd/zcwj_1940/bmjqtwj_1943/202403/t20240321_79787.shtml	
北京经济技术开发区关于加快打造 AI 原生产业创新高地的若干政策	2024.11.26	北京经济技术开发区管理委员会	https://www.beijing.gov.cn/gate/big5/www.beijing.gov.cn/zhengce/zhengcefagui/202503/t20250303_4024044.html	二（二）第三条
丰台区支持人工智能科技创新和产业创新融合发展的若干措施（征求意见稿）	意见征集	丰台区科学技术和信息化局	https://www.beijing.gov.cn/hudong/gfxwjzj/qjzjxx/202509/t20250918_4204925.html	附件 1 第二条
关于发布支持北京城市副中心数字经济高质量发展的实施指南（第一批）的通知	2025.5.9	北京市通州区经济和信息化局	https://www.bjtzh.gov.cn/bjtz/xxfb/202505/1753272.shtml	一（一）
怀柔区促进先进制造业及软件和信息技术服务业高质量发展扶持办法	2025.7.10	北京市怀柔区经济和信息化局	https://www.bjhr.gov.cn/zwgk/zcwj/202507/t20250715_4149752.html	第十二条

（二）广东省

广州海珠区：大模型企业首次完成国家级生成式人工智能（大语

言模型）上线备案的，给予最高100万元一次性奖励。

广州天河区：对上一年度完成国家级生成式人工智能服务备案的人工智能行业大模型+企业，按照上一年度研发投入的10%给予支持，每家企业每年支持最高不超过100万元；对上一年度完成国家级互联网信息服务算法或深度合成服务算法备案的人工智能行业大模型+企业，按照上一年度研发投入的10%给予支持，每家企业每年支持最高不超过20万元。

广州黄浦区：对首次完成国家级生成式人工智能（大模型）上线备案且模型落地应用场景具有示范性的企业，给予最高100万元一次性奖励。对首次完成国家级境内深度合成服务算法备案且具备良好应用的企业，给予最高20万元一次性奖励。

广州南沙区：鼓励人工智能企业备案大模型，对自主研发并已在国家互联网信息办公室生成式人工智能服务备案的通用大模型或者垂直领域大模型的所有权主体，按参数规模、模态数量等指标，给予最高100万元一次性奖励。

深圳福田区：对符合条件的企业购买已完成国家生成式人工智能服务备案的大模型进行应用场景开发和推广的，依条件按照不超过服务采购合同额（含模型私有化部署费用、云化部署费用、API调用费用以及大模型一体机中大模型部分费用）的30%给予支持，每家企业每年最高50万元。

深圳龙华区：对上年度成功通过国家互联网信息办公室《生成式人工智能服务备案》的模型，给予每个模型50万元，每家单位每年最高200万元的奖励

深圳龙岗区：对通过国家生成式人工智能服务备案登记的模型，

按每个模型最高50万元给予企业支持,同一企业每年最高不超过200万元支持。

深圳南山区：对符合条件的互联网或软件企业，算法纳入国家互联网信息办公室《境内深度合成服务算法备案清单》的，给予每个算法不超过100万，每家企业每年最高不超过500万元奖励。

珠海市：对通过国家深度合成服务算法备案的算法，给予每个不超过2万元，每家企业累计最高5万元奖励。对通过国家生成式人工智能服务备案的大模型，给予每个不超过25万元，每家企业累计最高50万元奖励。

东莞市：设立最高1500万元模型券，支持企业利用通过备案的大模型底座框架进行二次开发和应用推广，符合条件的给予最高100万元资助。

表 3 广东省生成式人工智能备案相关政策文件

文件名称	施行时间	发布机构	文件获取链接	具体提及处
广州市海珠区建设人工智能大模型应用示范区实施细则（征求意见稿）	2024.5.13（有效期3年）	广州市琶洲管委会	https://www.haizhu.gov.cn/hdjlpt/yjzj/answer/36270	附件1第三条（一）
广州市天河区促进人工智能行业大模型产业高质量发展的若干政策措施	2024.11.6（有效期3年）	广州市天河区科技工业和信息化局	https://www.gz.gov.cn/gfxwj/qjgfxwj/thq/qbm/content/mpost_9961289.html	第五条
广州开发区黄埔区支持人工智能产业高质量发展若干政策措施	2025.6.27（有效期3年）	广州市黄埔区工业和信息化局	https://www.gz.gov.cn/gfxwj/qjgfxwj/hpq/qbm/content/mpost_10331074.html	第三条
广州市南沙区促进	2025.7.1（有	广州市南沙	https://www.gz.gov	二、第八条

文件名称	施行时间	发布机构	文件获取链接	具体提及处
人工智能产业高质量发展扶持办法	效期至2027.12.31）	区人民政府办公室	.cn/gfxwj/qjgfxwj/nsq/qfb/content/mpost_10347095.html	
深圳市福田区支持人工智能产业高质量发展若干措施	2024.11.1（有效期至2027.12.31）	深圳市福田区科技和工业信息化局	http://www.szft.gov.cn/bmxx/qkjj/zcfg/content/post_11690882.html	第一条（二）
深圳市龙华区推动人工智能及机器人产业发展若干措施	2025.1.1（有效期3年）	深圳市龙华区科技创新局	http://www.szlhq.gov.cn/xxgk/zcfg/qgfxwj/qgfxwj_129575/content/post_11839311.html	第九条
深圳市龙岗区关于支持人工智能产业引领高质量发展若干措施	2025.1.10（有效期三年）	深圳市龙岗区工业和信息化局	https://www.lg.gov.cn/xxgk/zwgk/flfg/bmgfxwj/content/post_11932611.html	三（五）
南山区促进数字经济高质量发展专项扶持措施	2024.4.8（有效期三年）	深圳市南山区人民政府	https://www.szns.gov.cn/xxgk/bmxxgk/qkcj/xxgk/zcfg/zcfgjgfxwj/content/post_11229384.html	第四章第十三条
珠海市推动人工智能与机器人产业高质量发展若干措施	2025.5.23（有效期至2027.12.31）	珠海市工业和信息化局	https://www.zhuhai.gov.cn/zw/fggw/gfxwj/bmgfxwj/content/post_3800344.html	二
关于加快推动人工智能赋能制造业高质量发展的若干措施	2025.2.5（有效期3年）	东莞市人民政府	http://www.dg.gov.cn/zwgk/zfgb/szfwj/content/post_4377518.html	三

（三）上海市

上海市：对由大模型驱动的具有舆论属性或社会动员能力的互联网信息服务，开展常态化联系服务，加强合规指导，推动相关主体按照规定履行安全评估、算法备案等相关程序。在相关集聚区内，探索创新监管机制。

徐汇区：鼓励经营主体取得生成式人工智能服务登记、备案，经认定，可给予单个经营主体累计最高500万元奖励。对认定为上海市“专精特新”的企业，经认定，可给予最高10万元的一次性奖励。对认定为国家级专精特新“小巨人”的企业，经认定，可给予最高50万元的一次性奖励。

表 4 上海市生成式人工智能备案相关政策文件

文件名称	施行时间	发布机构	文件获取链接	具体提及处
上海市推动人工智能大模型创新发展若干措施（2023-2025年）	2023.10.20	市经济信息化委、市发展改革委、市科委、市委网信办、市财政局	https://www.shanghai.gov.cn/gwk/search/content/eb13ebcc74154857a6d42e48c6340917	四、10
徐汇区关于推动人工智能产业高质量发展的若干意见	2025.5.31（试行期二年）	徐汇区新型工业化推进办公室	https://www.shanghai.gov.cn/kjcx-gqwj4/20251126/8ca7a3e6b61b4db0a391b6fd0b5fe57b.html	二（十一）

（四）浙江省

浙江省：鼓励企业自主研发模型向中央网信办申请生成式人工智能模型备案，鼓励有条件的地方对通过备案的企业，依据其模型评测

等相关费用标准，给予一定政策支持。

杭州市：鼓励企业自研模型申请模型备案，对获得中央网信办生成式人工智能模型备案的企业，依据其模型评测等相关费用，给予不超过50万元的一次性奖励。

宁波市：关于开展《关于加快发展新质生产力 全力推进新型工业化的若干意见实施细则》政策申报的通知，自2023年1月1日以来对宁波市行政区域内依法注册登记、具有独立法人资格的企业或机构，通过国家大模型备案登记的企业或机构给与国家模型备案奖励。

嘉兴市：对获得中央网信办生成式人工智能模型备案的企业，给予100万元一次性奖励。对新入选国家、浙江省人工智能应用场景的企业分别给予50万元、20万元一次性奖励。

金华市：鼓励模型合规备案。鼓励企业自研模型申请模型备案，对获得中央网信办生成式人工智能模型备案的企业，依据其模型评测等相关费用标准，给予不超过50万元的一次性奖励。

嘉兴市平湖区：鼓励人工智能模型开发和推广应用。鼓励企业自研模型申请模型备案，对获得中央网信办生成式人工智能模型备案的企业，给予最高100万元一次性奖励。对新入选国家、浙江省人工智能应用场景、人工智能赋能新型工业化典型应用案例的企业分别给予最高50万元、20万元一次性奖励。

杭州市西湖区：支持模型合规备案。鼓励企业自研模型申请模型备案，对获得中央网信办、浙江省网信办生成式人工智能模型备案的企业，分别给予不超过50万元、10万元的一次性奖励。单个主体年度累计兑付该条款金额不超过100万元。

杭州市上城区：对获得中央网信办生成式人工智能模型备案的企

业，按照市级认定的奖励金额给予最高20万元的额外奖励。

杭州市滨江区：对采购并依托经备案的深度合成服务算法或生成式人工智能模型开展模型服务、智能体开发应用达到一定规模、具有良好成效的企业，经评审，按不超过上一年度交易结算费用的30%，给予单家企业最高200万元资助。

杭州市萧山区：对依托经备案的深度合成服务算法或生成式人工智能模型开展模型服务、智能体开发应用达到一定规模、具有良好成效的企业，经评审，每年择优评选不超过10个在萧山区成功落地并实现市场化的应用，给予300万元一次性奖励；其中，对提供基础大模型服务的，提高至800万元一次性奖励，同一应用不得重复申报。鼓励企业自研模型申请模型备案，对获得中央网信办生成式人工智能模型备案的企业，按市级政策给予配套支持。

宁波鄞州区：对开展模型训练与推理的人工智能企业，分别按经认定的模型应用、经备案的垂类模型，给予最高30、50万元的补助。

表 5 浙江省生成式人工智能备案相关政策文件

文件名称	施行时间	发布机构	文件获取链接	具体提及处
关于支持人工智能创新发展的若干措施	2025.5.31（有效期至2027.12.31）	浙江省政府	https://jxt.zj.gov.cn/art/2025/5/20/art_1229893033_2568658.html	六（二十）
支持人工智能全产业链高质量发展的若干措施	2025.8.19（有效期至2027.12.31）	杭州市人民政府办公厅	http://zfgb.hangzhou.gov.cn/mobile_detail.shtml?doc_id=5189685251.pdf	二（四）
关于加快发展新质生产力 全力推进新型工业化的若干意	2024.8.27	宁波市数字经济局	http://jxj.ningbo.gov.cn/art/2024/8/27/art_1229561613_58940308.html	七

文件名称	施行时间	发布机构	文件获取链接	具体提及处
见实施细则				
嘉兴市推动人工智能高质量发展实施方案	2024.10.10（有效期至2027年12月31日）	嘉兴市人民政府办公室	https://www.jiaxing.gov.cn/art/2024/12/5/art_1229782153_2072.html	三（七）
金华市关于支持人工智能创新发展的若干政策	2025.10.20（有效期3年）	金华市人民政府	https://www.jinhua.gov.cn/art/2025/8/25/art_1229562305_1824008.html	三（六）
平湖市人工智能新时代高质量发展实施方案	2025.5.15（有效期至2027.12.31）	平湖市人民政府办公室	http://www.pinghu.gov.cn/art/2025/4/16/art_1229456586_2549567.html	三（四）
西湖区进一步推动人工智能产业发展的若干措施	2025.4.13（有效期至2028.4.12）	西湖区经济和信息化局	https://zjjcmspublicnew.oss-cn-hangzhou-zwynet-d01-a.internet.cloud.zj.gov.cn/cms_files/jcms1/web1838/site/attach/0/26a8eb8a952e4322bef0f12b1338a7d8.pdf?fileName=26a8eb8a952e4322bef0f12b1338a7d8.pdf	二（二）4
上城区支持人工智能产业创新发展若干政策措施	2025.9.17（试行期1年）	杭州市上成区人民政府办公室	https://www.hzsc.gov.cn/art/2025/8/25/art_1229352497_1858953.html	二、2
关于加快新一代人工智能产业应用发展的若干意见	2025.6.10（有效期3年）	杭州市滨江区人民政府	http://www.hhtz.gov.cn/art/2025/5/9/art_1229225928_1851051.html	二（八）
关于促进人工	2025.8.15（有	杭州市萧山	https://www.xiaoshan.go	二（一）2

文件名称	施行时间	发布机构	文件获取链接	具体提及处
智能产业创新发展的若干政策（2025版）	效期1年）	区人民政府	v.cn/art/2025/7/14/art_1229293108_1856641.html	
鄞州区促进人工智能创新发展的若干政策措施（征求意见稿）	征求意见	宁波鄞州区经信局	https://www.nbyz.gov.cn/col/col1229879109/art/2025/art_4ad60d013bc74151aa44cd56c1d320da.html	一（三）

（五）四川省

成都市：对取得国家科技重大专项（含科技创新2030—重大项目）、国家重点研发计划立项项目成果或国家科学技术奖获奖成果并在蓉落地转化的，经评审择优，给予最高1000万元经费支持。支持企业、高校院所开展行业大模型研发应用，对性能先进且成功通过国家大模型备案登记的前十名，给予100万元一次性奖励。

成都高新区：支持开展人工智能算法备案。支持企业开展国家人工智能算法备案，企业完成中央网信办生成式人工智能服务备案的，给予50万元奖励;完成中央网信办深度合成服务算法备案的，给予5万元奖励。

绵阳市：对在绵企事业单位通过国家网信办生成式人工智能服务备案的大模型，按照大模型研发投入的10%，最高给予200万元奖励。

南充市：对通过国、省人工智能登记备案的大模型，按不超过实际研发投入10%的比例、给予最高不超过200万元一次性后补助。

德阳市：支持企业开展行业大模型研发，对通过国家网信办生成式人工智能服务备案的大模型，给予企业100万元奖励。

表 6 四川省生成式人工智能备案相关政策文件

文件名称	施行时间	发布机构	文件获取链接	具体提及处
成都市进一步促进人工智能产业高质量发展的若干政策措施	2024.2.17（有效期3年）	成都市经济和信息化局	https://cdjx.chengdu.gov.cn/cdsjxw/c160796/2024-01/22/content_e7ea3f6709434c64afe7c32ec0170066.shtml	第一章第二条
成都高新技术产业开发区关于支持人工智能（机器人）产业高质量发展的若干政策	2025.4.28（有效期3年）	成都高新区管委会	https://www.cdht.gov.cn/cdht/c139592/2025-07/11/b205bdad62154360a44537d7fd917cc1/files/23154f92014b4b4992b93cfada05b9ff.pdf	三（五）
绵阳市支持人工智能产业发展若干政策	2024.6.14（试行期1年）	绵阳市人民政府	https://www.my.gov.cn/mysrmzf/c100061/202406/dbbf87d486d54c49b089d09d3d909f77.shtml	第二条
南充市支持人工智能产业高质量发展的若干措施	2025.4.3（有效期3年）	南充市人民政府	https://www.nanchong.gov.cn/zwgk/zfgb/2025n/d3q/202504/t20250425_2148715.html	二
德阳市支持数字经济高质量发展若干措施	2025.7.10（有效期3年）	德阳市人民政府	https://credit.deyang.gov.cn/snzc/20250708/27677.html	七

（六）江苏省

苏州市：加大对国家人工智能重大工程和应用场景布局支持力度，对牵头申报主体、联合申报主体、合作申报主体，分别给予最高1亿

元、2000 万元、500 万元支持。

无锡市：对首次完成国家级生成式人工智能模型备案的企业，给予最高 50 万元的一次性奖励，再按模型应用落地企业数量分档给予每个最高 30 万元的累加奖励，单个企业获得上述奖励合计最高 200 万元；对首次完成国家级境内互联网信息服务算法备案或深度合成服务算法备案的企业，给予最高 5 万元的一次性奖励。

南京玄武区：对通过国家网信办《生成式人工智能服务管理暂行办法》备案的企业、高校、科研院所等各类企事业单位，给予 20 万元一次性区级奖励。

表 7 江苏省生成式人工智能备案相关政策文件

文件名称	施行时间	发布机构	文件获取链接	具体提及处
苏州市高水平建设“人工智能+”创新发展试验区的若干措施	2025.3.17（有效期至2028.3.16）	苏州市人民政府	https://www.suzhou.gov.cn/szsrmzf/zfwj/202502/9123c26e54774cfb8cf5a2a336460ca2.shtml	二（五）
市政府关于建设“人工智能+”标杆城市的政策意见	2025.5.23（有效期至2027.12.31）	无锡市人民政府	https://www.wuxi.gov.cn/doc/2025/05/19/4576497.shtml	六
玄武区关于促进大模型产业发展的若干措施	2025.11.10（有效期1年）	南京市玄武区人民政府	http://www.xwzf.gov.cn/xwqrmzf/202510/t20251014_5666740.html	三

（七）安徽省

合肥市：对获得国家生成式人工智能服务备案的中国声谷企业，给予最高 100 万元奖励。

表 8 安徽省生成式人工智能备案相关政策文件

文件名称	施行时间	发布机构	文件获取链接	具体提及处
合肥市推动中国声谷建设和人工智能产业高质量发展若干政策	2025.1.1（有效期3年）	\	http://www.hfusp.com/index/Article/info.html?cate=92&id=1099	六

（八）湖北省

武汉市：支持企事业单位开展人工智能大模型研发、备案和落地应用。围绕电子信息制造、工业质检、教育、医疗、遥感、文创、金融等领域每年遴选一批性能先进的垂直行业模型，对牵头研发单位按照研发成本的30%给予最高500万元补助，并对其使用算力费用给予每年50%、最长三年、累计最高500万元补助。支持在全市各行业、各领域开展人工智能大模型先行先试应用。

武汉市洪山区：鼓励企业自主研发模型向国家互联网信息办公室申请生成式人工智能模型备案，对通过备案的企业，按每个模型100万元给予企业支持，同一企业每年最高不超过200万元。

表 9 湖北省生成式人工智能备案相关政策文件

文件名称	施行时间	发布机构	文件获取链接	具体提及处
武汉市促进人工智能产业发展若干政策措施	2025.3.19（有效期2年）	武汉市人民政府	https://3g.wuhan.gov.cn/zwgk/xxgk/zfwj/gfxwj/202502/t20250218_2535438.shtml	三
洪山区支持人工智能产业高质量发展若干措施	2025.9.12（有效期至2027.12.31）	武汉市洪山区人民政府	https://www.hongshan.gov.cn/xxgk/zc/qzfgfxwj/202508/t20250812_2632464.shtml	六

（九）福建省

厦门市：推广模型服务。设立“模型券”，对企业购买通过国家网信办备案的模型开展智能体研发、提供模型服务的，按照模型购买费用给予最高不超过30%的补助，每家每年最高不超过200万元。

厦门集美区：获得算法备案的，一次性给予10万元奖励。对调用符合条件的大模型服务商API的，按当年度实际使用金额35%给予奖励，最高不超过30万。

厦门思明区：对2024年1月1日（含）以来，模型开发企业首次完成国家级境内互联网信息服务算法备案或深度合成服务算法备案、首次完成国家级生成式人工智能服务备案，年度实际服务5家（含）以上非关联企业，且该算法或模型产品已上架到AI产业服务平台上的，分别给予5万元、20万元一次性科技研发奖励。

表 10 福建省生成式人工智能备案相关政策文件

文件名称	施行时间	发布机构	文件获取链接	具体提及处
厦门市进一步推动人工智能产业发展若干措施	2025.4.21（有效期至2027.12.31）	厦门市人民政府办公厅	http://www.xm.gov.cn/zfgb/99833851	三（九）
集美区产业高质量发展若干措施（暂行）（2025版）	2025.1.24（有效期至2025.12.31）	厦门市集美区人民政府	http://www.jimei.gov.cn/xxgk/xxgk/fgwj/yxwj/202501/t20250124_1087797.htm	四、第九条
思明区促进新一代人工智能产业发展若干措施	2025.1.1（有效期至2026.12.31）	厦门市思明区科技和工信局、思明区财政局	http://www.siming.gov.cn/zfxxgkzl/qrmzf/gfxzc/xzgfxwj/202412/t20241220_1083000.htm	第五条

（十）山东省

青岛市：对参数量不低于千亿的通用大模型或参数量不低于百亿的行业大模型、典型应用场景不少于5个、按规定通过国家有关部门大模型备案登记的，按照大模型建设方实际支付算力费用不超过20%的比例，给予最高500万元的补贴。

表 11 山东省生成式人工智能备案相关政策文件

文件名称	施行时间	发布机构	文件获取链接	具体提及处
青岛市2024年“促进经济巩固向好、加快绿色低碳高质量发展”政策清单（第二批）	2024.7.5	青岛市人民政府	http://www.qingdao.gov.cn/zwgk/zdgk/fgwj/zcwj/szfgw/gw_2024/qzf/202407/t20240710_8120595.shtml	一、11

（十一）贵州省

贵阳市：对通过国家网信办境内深度合成服务算法备案的大模型，每年择优选择一批应用效果好、产业贡献高的大模型，按其建设成本的30%给予一次性补助，最高不超过500万元。

表 12 贵州省生成式人工智能备案相关政策文件

文件名称	施行时间	发布机构	文件获取链接	具体提及处
贵阳贵安关于支持人工智能大模型发展的若干措施（试行）	2024.1.19（有效期3年）	贵阳市大数据发展管理局	https://dsjj.guiyang.gov.cn/newsite/zwgk/zdlyxxgk/cyrhfz/202401/t20240125_83627561.html	二

（十二）广西省

南宁市：对获得国家互联网信息办公室生成式人工智能模型备案的企业，依据该模型技术评测（或相应评估）情况以及模型评测相关费用等情况，给予最高50万元奖励。

表 13 广西省生成式人工智能备案相关政策文件

文件名称	施行时间	发布机构	文件获取链接	具体提及处
南宁市支持中国—东盟人工智能创新合作中心高质量发展第一批政策措施的通知	2025.5.7（有效期至2027.12.31）	南宁市人民政府	https://www.nanning.gov.cn/zwgk/fdzdgknr/zfgb/2025nzfgb/d8q/zfwj/t6328104.html	三（十一）

（十三）宁夏回族自治区

银川市：凡是正式登记注册的企业或分支机构，近3年内未发生重大安全、重大质量事故或严重环境违法行为都可申报。其中，申报AI大模型项目需满足AI大模型需由大模型企业原创研发、训练，训练参数不少于100亿，已完成国家级生成式人工智能（大语言模型）上线备案或国家级境内互联网信息服务算法备案或深度合成服务算法备案，具有自主知识产权等条件；申报人工智能应用场景项目需能够解决行业痛点、提升服务质量、优化工作流程等特点。

表 14 宁夏回族自治区生成式人工智能备案相关政策文件

文件名称	施行时间	发布机构	文件获取链接	具体提及处
银川市启动支持AI大模型、人工智能应用场景项目申报工作	2024.7.3	银川市人民政府	https://www.yinchuan.gov.cn/xwzx/mrdt/202407/t20240703_4584342.html	\

（十四）山西省

大同市：对自主研发并在国家网信办完成生成式人工智能服务备案的通用大模型或者垂直领域大模型的所有权主体，额外给予100万元一次性奖励。

表 15 山西省生成式人工智能备案相关政策文件

文件名称	施行时间	发布机构	文件获取链接	具体提及处
大同市促进人工智能产业高质量发展若干措施	2025.9.18（有效期3年）	大同市政府	https://www.dt.gov.cn/dtszf/zxgb/202509/cbced2b473e74833b1eba1af7226e874.shtml	三

三、生成式人工智能服务备案现状

随着人工智能技术的迅速发展，生成式人工智能服务的应用领域愈加广泛。与此同时，大模型引发的安全和合规风险亦成为社会关注焦点。为促进生成式人工智能服务创新发展和规范应用，国家互联网信息办公室会同有关部门发布《生成式人工智能服务管理暂行办法》。该办法第十七条明确规定提供具有舆论属性或者社会动员能力的生成式人工智能服务的，应当按照国家有关规定开展安全评估，并按照《互联网信息服务算法推荐管理规定》履行算法备案和变更、注销备案手续。

对生成式人工智能服务提供者来讲，完成生成式人工智能服务备案是其法定义务，亦可推动其建立健全安全管理制度和技术措施，提升生成式人工智能服务的安全性和可靠性，提高用户信任度；同时，许多政府将完成生成式人工智能服务备案作为企业享受扶持政策的前提条件，为企业提供资金支持。对大模型服务使用者而言，备案信息公示也有助于用户了解大模型的运行基础和使用风险，保障自身的

合法权益。

自人工智能浪潮兴起以来，国内生成式人工智能服务发展迅猛，“十四五”时期，我国已有超 700 款生成式人工智能服务完成备案。[1]同时，从国家网信办发布的阶段性公告看，截至 2025 年 11 月 1 日，全国累计 611 款生成式人工智能服务完成备案；仅 2025 年 9—10 月就新增 73 款完成备案，并新增登记 35 款通过 API 等方式调用已备案模型能力的应用或功能（累计登记 306 款）。从该阶段公告附件的新增清单统计看，新增备案在部分地区相对集中，例如广东、上海在该阶段各新增 14 款。备案与登记主体覆盖科技与互联网企业、高校科研机构、央国企及创新型企业等，相关服务形态既包括通用能力，也包括面向行业垂类与多模态场景的应用落地。

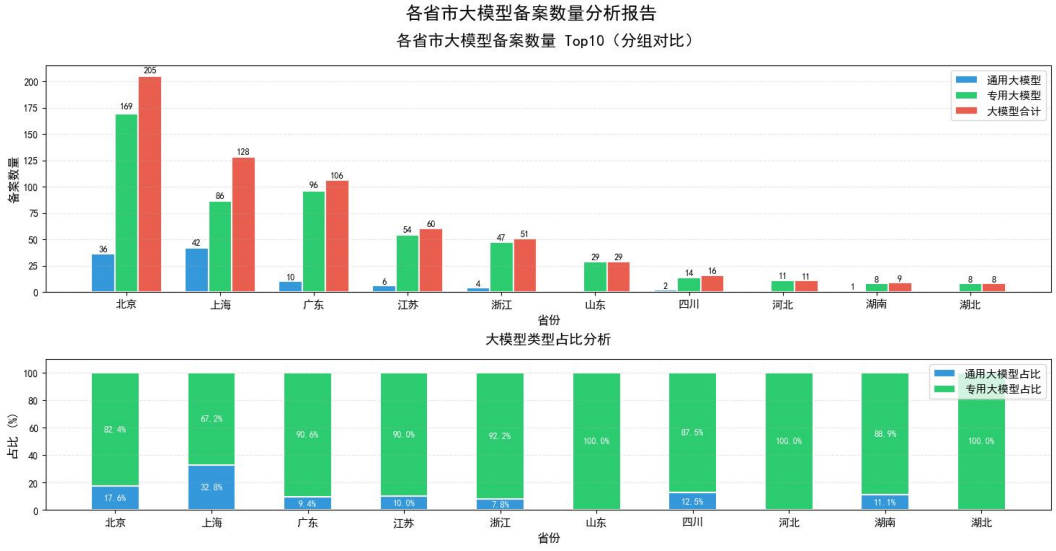


图 3 各省市大模型备案数量 TOP10

[1] 参见《超 700 款生成式人工智能大模型产品完成备案》，https://www.xinhuanet.com/tech/20251225/cb5139a090fe404d84fb6897e46b9f70/c.html，最后访问日期:2025 年 12 月 27 日。

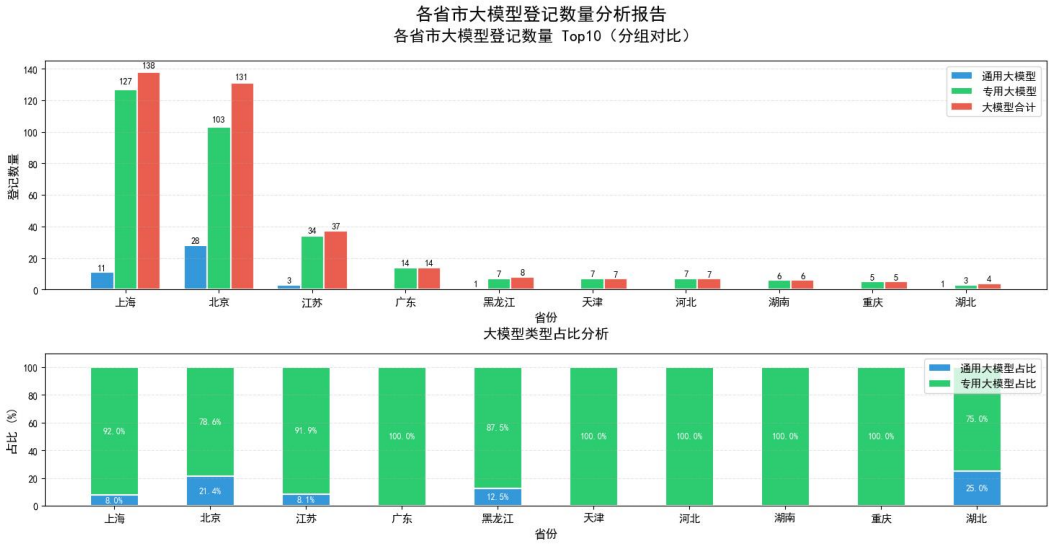


图 4 各省市大模型登记数量 TOP10

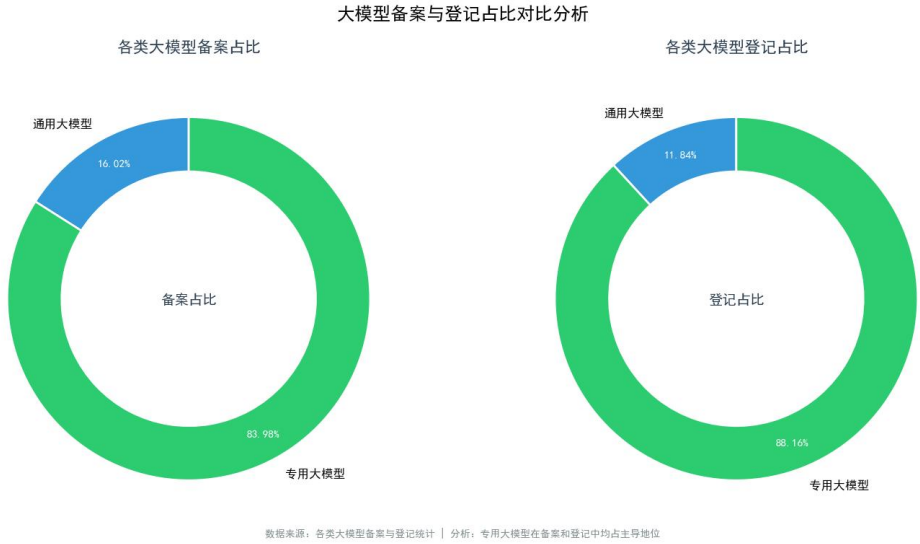


图 5 大模型备案与登记占比对比

值得注意的是，根据国家网信办公告，北京市的奔驰虚拟助手大模型，上海市的特斯拉 xBot 客户服务、沃尔沃小沃智能助手，成为首批通过备案的外企产品，这标志着我国大模型产业在国际化发展中迈向新阶段。另北京日报 2025 年 12 月 1 日发文称北京累计备案上线生成式人工智能服务 183 款，数量全国第一。豆包、智谱等标杆模型性能稳居全球第一梯队。2025 年 12 月 2 日中央网信办主任庄荣文答

记者问时亦表示，截至 2025 年 12 月 1 日，我国已有 663 款生成式人工智能服务完成备案。[2]与 2025 年 11 月 1 日的 611 款相比，新增 52 款生成式人工智能服务完成备案，较之上期数据增长了 8.51%。可见，国内生成式人工智能服务备案已是人工智能行业发展的重要环节，既有利于推动技术创新进步和应用落地实施，又有助于保障行业持续健康发展。随着政策的进一步完善和技术的成熟，生成式人工智能服务备案数量有望继续增长。

[2] 参见《推动网信事业高质量发展开创网络强国建设新局面》，载 https://www.cac.gov.cn/2025-12/02/c_1766396564941361.htm，最后访问日期:2025 年 12 月 27 日。

第三章 生成式人工智能服务备案难点与流程解析

一、生成式人工智能服务的备案要求

生成式人工智能服务，根据《生成式人工智能服务管理暂行办法》第二条规定，系指利用生成式人工智能技术向中华人民共和国境内公众提供生成文本、图片、音频、视频等内容的服务。本办法第十七条规定，提供具有舆论属性或者社会动员能力的生成式人工智能服务的，应当按照国家有关规定开展安全评估，并按照《互联网信息服务算法推荐管理规定》履行算法备案和变更、注销备案手续。《生成式人工智能服务管理暂行办法》明确提供者的义务与责任，提供者是利用生成式人工智能技术提供生成式人工智能服务的组织、个人，包括通过提供可编程接口等方式提供生成式人工智能服务的组织、个人。[3]由此可知，企业自主研发的生成式人工智能服务，若面向境内公众提供生成内容服务，且该服务具有舆论属性或社会动员能力，则需进行备案。即使调用第三方已备案的大模型，但经过二次开发或微调，且调整后的模型面向公众提供具有舆论属性或社会动员能力的服务，同样需要备案。

二、生成式人工智能服务的备案难点

生成式人工智能服务备案侧重点在于模型的可靠性、可追溯性和合规性，要全面审查模型架构、数据来源、应用场景等方面，确保生成内容合乎伦理，数据来源合法可追溯。

申请生成式人工智能服务备案需以安全、合规为中心，确保提供的材料真实可靠，内容完整、逻辑严谨。核心材料包括：《生成式人工智能服务上线备案申请表》《安全评估报告》《模型服务协议》《语

[3] 参见《专家解读丨推动生成式人工智能精细化治理》。载 https://www.cac.gov.cn/2023-07/13/c_1690898363806525.htm，最后访问日期:2025 年 12 月 27 日。

料标注规则》、《拦截关键词列表》《评估测试题集》，具体材料各地方网信办要求可能略有不同。

（一）生成式人工智能服务上线备案申请表

生成式人工智能服务上线备案申请表是备案的核心申请文件，需准确、详尽填写，各项信息应与其他材料保持一致。主要包括：基本情况、模型研制、服务与安全防范、安全评估、附件及备注等部分。

基本情况需填写模型名称、单位名称、负责人、联系电话、单位属地、主要功能（分为人机对话、文字生成、图片生成、声音合成、视频合成、代码生成或优化六类，根据自己模型的基本情况按需勾选）、适用人群、适用场合（自动控制、医疗信息服务、心理咨询、关键信息基础设施、其他）、限定领域（未限定领域即通用领域；限定领域主要针对于医疗、金融、教育等行业的垂类大模型）

模型研制需说明模型备案情况，是否采用已备案模型，以及训练语料和标注语料来源与规模。

服务与安全防范含推理算力资源和服务等内容，包括开发服务能力，采用第三方云平台或服务器情况，服务方式，为保障数据安全、隐私保护和内容合规性所采取的具体技术和管理措施等项。

安全评估需填写基本情况和评估情况等内容，如语料内容评估、生成内容评估、涉知识产权及商业秘密评估、涉民族信仰性别等评估、涉安全性有效性可靠性评估、模型性能（拒答率）评估。

附件为安全评估报告、模型服务协议、语料标注规则、拦截关键词列表、评估测试题集。

（二）安全评估报告

生成式人工智能服务备案的安全评估是整个流程的关键环节，内

容需详实且可追溯。安全评估应依据《生成式人工智能服务安全基本要求》的要求，确保数据来源安全，数据内容管理规范，数据标注安全，符合模型安全要求，采取相应安全措施，防范训练数据及生成内容的五大主要安全风险。

具体要求如下：

一是语料安全评估，涉及文本训练语料规模、各类型语料规模、训练数据来源（开源、自采、商业语料）、语料标注数量、标注人员情况、标注规则、标注内容准确性核验、语料合法性。

二是模型安全评估，包括语料内容评估、生成内容评估、涉知识产权及商业秘密评估、涉民族信仰性别等评估、涉安全性有效性可靠性评估、模型性能（拒答率）评估。

三是安全措施评估，如模型适用人群、场合、用途；服务过程中收集保存个人信息情况；个人信息知情状况；图片、视频标注情况；投诉举报情况；非法内容拦截措施；模型更新、升级措施。

四是总体结论，安全评估报告应表明即将上线的生成式人工智能服务是否合乎《生成式人工智能服务管理暂行办法》规定，存在的主要风险，风险预判及防范措施以及生成式人工智能服务提供者面对风险的处置能力。

（三）模型服务协议

模型服务协议系规范生成式人工智能服务提供者与使用者之间双方权利义务的重要文件，涵盖待备案生成式人工智能服务当前提供的所有服务内容，并适用于未来可能推出的任何新功能、新产品或服务，以确保服务的统一性与完整性。协议内容需合法合规，充分保护用户权益，并清晰界定各方责任。

协议应表明除非适用法律法规另有规定，所有新增功能或服务均受本协议条款的约束。本协议可规定账号使用注册说明、账户安全保障与责任、账户转让规则与限制、更新与维护指南、服务提供者权利与义务概述、服务使用者权利和义务、协议终止条件及程序、协议终止后的处理等内容。注意需明确用户隐私条款，包括用户数据的收集、使用、存储、共享等方面的规定。提供用户投诉举报的渠道和处理流程，保障用户合法权益。明确在特定情况下的开发者责任限制和免责范围。包含产品及服务的各项规则及隐私条款等，需协同法务共同制定提交，明确服务范围、双方权利义务、数据使用与保护、违约责任等。

（四）语料标注规则

语料标注规则主要用于数据标注、模型训练和生成内容的审核。语料标注应遵循合法合规、价值观导向、尊重合法权益、公平包容、透明准确的原则。符合《中华人民共和国网络安全法》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》《生成式人工智能服务管理暂行办法》《人工智能生成合成内容标识办法》《生成式人工智能服务安全基本要求》等法律法规的要求。标注流程、标注规则的实施细则、标注示例等内容亦应有所体现。

就标注流程而言，内容审核是标注流程的开端，可利用关键词库、NLP 检测等工具，通过预筛选、人工审核、风险评级等方式，识别并过滤不合规或敏感内容。数据分类和标注是构建高质量数据集的关键步骤。根据数据内容的性质，将数据进行分类，设置详细的标注标签。同时应特别关注数据中潜在的偏见与歧视问题，可采用自动识别、复核等方式在标注过程中剔除含偏见内容的数据，并对可能影响生成结

果的偏见内容，进行权重调整或数据补充。知识产权保护和伦理合规是标注工作的重要环节，应检查数据来源是否合法合规，是否取得完整授权以及是否侵犯第三方的商标、专利、版权等，标注内容是否合理引用了开放许可的数据，数据是否涉及伦理争议等。标注的最终目标是确保生成内容具有高准确性和透明性，应检查标注内容是否具备事实依据，对标注数据的来源、处理方法进行详细记录。多维度进行评估，丰富评估方式，通过质量评估提升标注的整体规范性，对标注过程中出现的错误，定期反馈并组织优化。

标注规则的实施细则是保障数据合法性、合规性和高质量的关键环节。标注团队的组织架构、人员管理、工具支持、执行步骤、质量控制、反馈机制等方面均至关重要。建立清晰的责任分工，确保标注规则执行到位。组织标注人员培训，确保标注人员具有足够的能力和合规意识。优化标注工具及工具使用流程，通过工具提高标注效率，降低人工工作负担。将规则落实到标注操作的每一步，确保执行一致性。设置质量检查频率、质量检查指标通过质量监控机制，确保标注结果符合预期。完善反馈流程，定期更新标注规则，通过反馈闭环，不断优化标注规则和流程。模拟应急场景，制定应急措施，快速响应标注过程中的重大问题。

（五）拦截关键词列表

拦截关键词列表应在合法合规的前提下尽可能多的扩大规模，选择建议至少包含 10,000 个关键词，个别要求严格的省市关键词数量会更高，另关键词需覆盖多种语言和新出现的风险词汇（如涵盖政治、暴力、违法等各方面内容）。关键词拦截也应精准定义并定期更新，

确保其时效性和有效性，避免错误拦截和遗漏。考虑到不同文化背景下语言表达的差异，关键词列表建议含多种语言和表达方式。

（六）评估测试题

评估测试题集需具有代表性和全面性，能够有效检验模型的安全防护能力和内容审核机制。测试结果（包括模型的回答和拒答情况）通常作为安全评估报告的重要支撑材料。生成内容测试题集需覆盖多个测试场景（参考《生成式人工智能服务安全基本要求》），验证模型拒答率与合规性。非拒答测试题库覆盖文化、历史、法律等多方面内容，确保生成内容符合核心价值观。

（七）其他辅助材料

根据不同地方网信办的要求，申请备案时可能需要提供模型训练数据来源及知识产权合规性说明；安全负责人、技术负责人任命文件及联系方式；网络安全等级保护备案证明；测试账号及使用说明等文件。

三、生成式人工智能服务的备案流程

生成式人工智能服务备案整体周期较长，需按阶段推进，备案主要涉及以下流程：

（一）备案前准备

前期准备阶段，认真研读相关政策和标准，如《中华人民共和国网络安全法》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》《生成式人工智能服务管理暂行办法》《生成式人工智能服务安全基本要求》《互联网信息服务算法备案管理规定》等文件及属地网信办具体要求，建议组建含技术、法务、合规人员在内的专

业队伍，全面梳理模型研发流程、语料合法性、安全机制等，确保合法依规。

结合下表并根据生成式人工智能服务的规模、服务范围和影响，确定是否可以备案。准备初步的备案申请材料，包括企业资质、生成式人工智能服务简要介绍、安全评估初步构想等。通过网信办提交备案意向，并确保所有信息真实准确。网信办审核后，会反馈需要提交的具体材料清单，并可能指派指导老师协助企业完成备案表的填写。企业需按照网信办要求认真填写备案表中的所有项目。

表 16 备案前期工作

数据处理		**模型训练**		**优化调整**		
预计周期：1个月	工作内容： (1)对现有数据进行系统化整理。 (2)根据数据重要性进行分级，清洗噪声数据，确保数据质量。 (3)批量构建无监督训练语料： (4)基于现有数据构建优质训练语料。 (5)设计拒答问题集。	预计周期：2个月	工作内容： (1)准备训练代码和环境，确保训练资源充足。 (2)开展模型训练。 (3)设定评估指标，评估模型性能并和优化。 (4)测试模型侧的拒答能力。 (5)模型部署。	预计周期：2个月	工作内容： (1)增加拦截策略，弥补模型缺陷进行模型微调。 (2)模型内测。	**提交备案**

（二）备案申请

生成式人工智能服务备案由网信部门通知或企业自行发起，企业根据网信办提供的备案表要求和评估要点，系统性地准备各项备案材

料。企业需组织内部力量进行严格的自评估和自测试，重点关注：应答/拒答的准确率、生成内容的合规率、关键词拦截的有效性、训练语料来源的合法性、各项安全保障措施是否健全到位等。确保所有提交的材料之间信息一致，逻辑清晰，避免出现矛盾之处。 仔细检查材料的格式规范和语言表达，力求专业、严谨、易于理解。在正式提交前进行最终确认，确保万无一失并准备一个或多个功能完善、可供监管部门随时访问和测试的产品账号。在完成所有材料的准备和测试账号的设置后，将全套材料正式提交给属地网信办进行初步审核。企业应向所在地省级（直辖市为市级）网信办线下提交装订成册的纸质材料和模拟测试账号。需注意所有纸质版材料均需加盖企业公章，电子版材料应与纸质材料保持完全一致。

属地省级网信办会重点审核材料的完整性、安全性、合规性以及模型的初步情况，同步开展技术安全测试。如果初审通过，材料将被上报至中央网信办进行复审。如果初审未通过，属地省级网信办会给出修改意见，企业需根据意见进行自查调整，并重新提交。若审核不通过，会下发补正通知，企业须在合理期限作出修改并重新提交。通过省级网信办初审则可至中央网信办复审，中央网信办将组织专家对备案材料和模型进行更为深入和全面的审查，征求工信、公安等意见并再次开展技术测试。复审通过后，中央网信办会在其官方渠道公示通过备案的企业名单和生成式人工智能服务信息，并下发生成式人工智能服务备案号。企业可在服务页面公示其备案号以供使用者查询。详见下图：

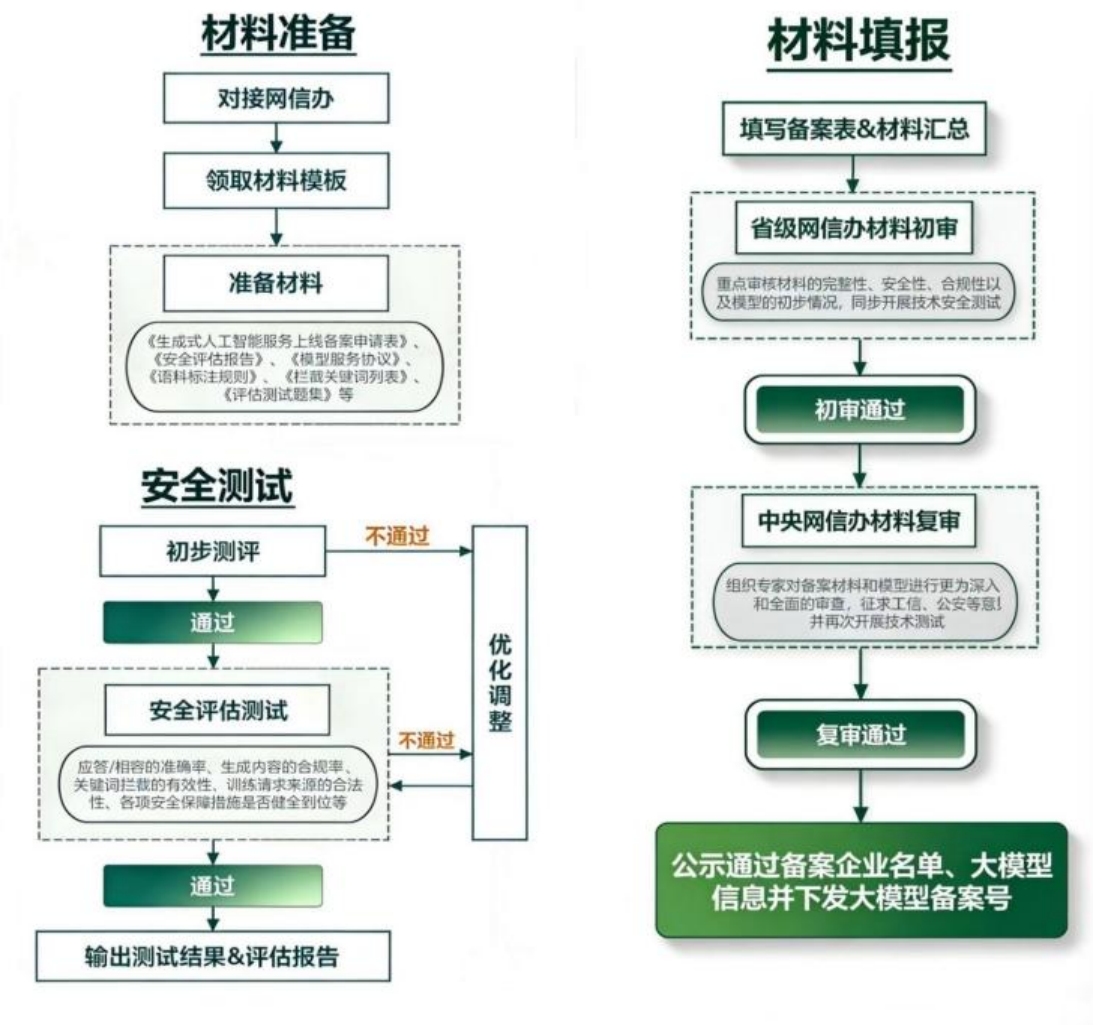


图 6 大模型备案流程图

（三）后续维护

备案通过不意味着可以一劳永逸，企业仍需定期维护生成式人工智能服务以满足合规要求。建议设置动态更新机制，定期更新拦截关键词库和测试题库，对照《生成式人工智能服务安全基本要求》《人工智能生成合成内容标识办法》开展内部评估，留存评估结果及问题整改记录。

四、生成式人工智能服务的登记流程

生成式人工智能服务登记是指网信部门按照《生成式人工智能服务管理暂行办法》及相关要求，对通过 API 接口或其他方式直接调用

已备案生成式人工智能服务能力，且面向境内公众提供具有舆论属性或者社会动员能力的生成式人工智能服务开展的管理方式。仅调用已备案生成式人工智能服务 API 的企业，未做任何调整的企业需依据《生成式人工智能服务管理暂行办法》开展生成式人工智能服务登记工作，**其不需要提交安全评估报告**。

所需材料主要为上线备案表（基本情况+模型调用+服务与安全防范+安全评估+自愿承诺+附件）；调用已备案生成式人工智能服务情况说明及相关证明材料；产品服务协议（含产品使用协议、隐私保护政策）；安全管理制度（管理制度中要包含“非法内容拦截标准”章节）；拦截关键词列表；评估测试题；测试通道（测试账号及 API 接口）；其他材料。

登记流程为：1、获取登记表：向属地网信办报备，获取“生成式人工智能服务登记表”等材料；2、准备上述材料，企业内部展开评估，编写相关材料，准备测试账号；3、提交审核：提交材料和测试账号给省级网信办审核；4、省级审核：省级网信部门进行材料审核及技术测试（安全测试）；审核通过后，上报中央网信办；如未通过，修改材料或调整模型能力后再次提交审核；通过后即可获得上线编号。详见下图：

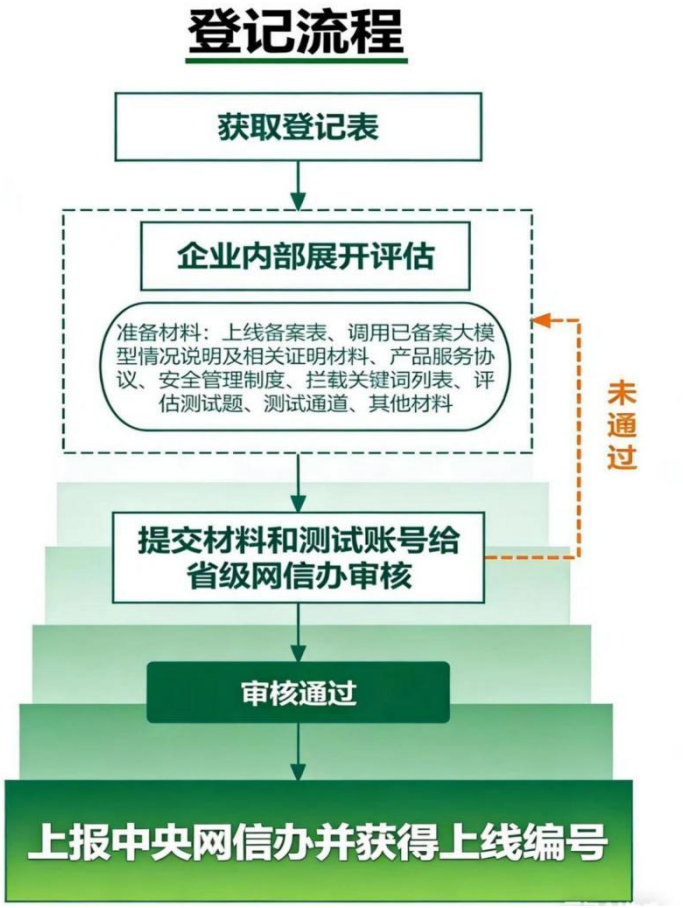

登记流程
获取登记表
企业内部展开评估
准备材料：上线备案表、调用已备案大模型情况说明及相关证明材料、产品服务协议、安全管理制度、拦截关键词列表、评估测试题、测试通道、其他材料
未通过
提交材料和测试账号给省级网信办审核
审核通过
上报中央网信办并获得上线编号

图 7 大模型登记流程图

第四章 生成式人工智能服务备案的合规风险

根据《生成式人工智能服务管理暂行办法》及相关法律法规，未依法履行备案义务的服务提供者应承担相应的法律责任。本部分将力求全面地梳理服务提供者未履行生成式人工智能服务备案义务的主要类型和法律责任，并针对服务提供者的常见错误进行实务经验总结，以期为生成式人工智能服务提供者提供具有针对性的合规指引。

一、未履行生成式人工智能服务备案义务的主要类型

根据《生成式人工智能服务管理暂行办法》和网信部门的具体要求，提供具有舆论属性或者社会动员能力的生成式人工智能服务的，可通过属地网信部门履行备案义务（通过 API 接口或其他方式直接调用已备案模型提供生成式人工智能服务仅需登记），这是生成式人工智能服务提供者核心的前置性合规义务。[4]在实践中，未履行生成式人工智能服务备案义务主要有以下几种常见类型：

1.“未备案”：指完全没有履行法定的备案义务便上线服务产品。在 2025 年 9 月 16 日国家网信办发布的网络安全、数据安全、个人信息保护相关执法典型案例之“浙江某科技有限责任公司运营的 App 提供深度合成服务未按规定进行安全评估案”便属于此种类型。案涉企业运营的 App 是一款深度合成类服务产品，提供视频换脸、图片换脸、照片舞动配音等图片处理功能，用户可对上传图片、视频中的人物进行换脸，但未按规定落实安全评估要求，存在较大安全风险。[5]实践中，还存在拒不改正，仍未备案重新上线的违法情形。上海市网信办在 2025 年的专项行动中就曾对 3 家拒不整改的生成式人工智能服务

[4] 参见《关于发布生成式人工智能服务已备案信息的公告（2025 年 9 月至 10 月）》，载 https://www.cac.gov.cn/2025-11/11/c_1764585284364412.htm，最后访问日期：2025 年 12 月 27 日。

[5] 参见《国家网信办发布近期网络安全、数据安全、个人信息保护相关执法典型案例》，载 https://www.cac.gov.cn/2025-09/16/c_1759741437315419.htm，最后访问日期：2025 年 12 月 19 日。

网站，首次适用《生成式人工智能服务管理暂行办法》予以立案处罚。[6]

2.“漏备案”：此种类型通常源于服务提供者的判断失误，该失误主要表现在服务提供者对服务是否“具有舆论属性或社会动员能力”，以及未能正确区分单纯的算法备案和人工智能模型备案之间的重大差异等方面。在法律后果上，“漏备案”与“未备案”性质类似，一旦被监管部门认定属于应备案而未备案，即需承担相应的法律责任。

3.“假备案”：指服务提供者通过提交虚假材料骗取备案，或备案后发生重大变更未履行变更备案手续。此行为性质更为恶劣，不仅构成对监管秩序的公然欺诈，也使备案状态形同虚设。除了要承担与“未备案”同等的法律责任外，还可能会被认定为“情节严重”。

二、未履行生成式人工智能服务备案义务的法律责任

《生成式人工智能服务管理暂行办法》第二十一条是关于服务提供者未履行生成式人工智能服务备案义务而应当承担法律责任的总括性条款。该条款明确违反本办法规定：

1.依据《网络安全法》《数据安全法》《个人信息保护法》《科学技术进步法》等上位法处罚；

2.上位法未规定的，由有关主管部门给予警告、通报批评、责令改正；拒不改正或情节严重的，可责令暂停服务；

3.该条还实现了与《治安管理处罚法》和《刑法》的衔接，规定构成违反治安管理或犯罪的，将依法追究相应责任。

根据公开网络检索，2025 年的典型监管案例有：

[6] 参见《上海市网信办整治 AI 滥用：下线 54 款违规 App，立案处罚 3 家网站》，载 https://www.thepaper.cn/newsDetail_forward_32030475，最后访问日期：2025 年 12 月 19 日。

表 17 2025 年典型监管案例

案例类型	地区	违法行为	法律责任
未进行安全评估与备案	浙江[7]（国家网信办公布）	某科技公司运营的 App 提供 AI 换脸服务，未按规定进行安全评估即上线。	应用程序被下架
未进行安全评估与备案	江苏[8]	累计排查提供文本生成、图片生成、音频生成等人工智能应用但未履行生成式人工智能服务备案（登记）的网站 50 个，APP 应用 42 个。	督导限期备案
“拒不改正”，违法重新上线	上海[9]	在“回头看”巡查中，发现 3 家已被要求下线功能的网站，在未通过安全评估的情况下，自行将生成式 AI 功能重新上线	上海市网信办首次适用《生成式人工智能服务管理暂行办法》予以立案处罚

因此，生成式人工智能服务的安全评估与备案是服务提供者的核心义务，这是监管的首要关注点。只要服务具有舆论属性或社会动员能力，上线前必须完成安全评估与备案。上述案例表明，未履行此义务是导致服务被下架和处罚的主要原因。此外，对于已发现的违法问题，如果企业拒不整改，监管部门会从责令改正升级为立案处罚。

[7] 参见《国家网信办发布近期网络安全、数据安全、个人信息保护相关执法典型案例》，载 https://www.cac.gov.cn/2025-09/16/c_1759741437315419.htm，最后访问日期：2025 年 12 月 27 日。

[8] 参见《江苏省委网信办深入开展“清朗·整治 AI 技术滥用”专项行动第一阶段工作》，载 http://jswx.gov.cn/zhengce/qinglang/202506/t20250613_108401.shtml，最后访问日期：2025 年 12 月 27 日。

[9] 参见《上海市网信办整治 AI 滥用：下线 54 款违规 App，立案处罚 3 家网站》，载 https://www.thepaper.cn/newsDetail_forward_32030475，最后访问日期：2025 年 12 月 27 日。

三、生成式人工智能服务企业的常见错误与经验总结

在《生成式人工智能服务管理暂行办法》的监管框架下，服务提供者面临的不仅是法律条文，更是一系列复杂的实践挑战。从服务上线前因疏忽或准备不足导致的备案失败，到运营中因机制缺失引发的安全事件，常见的错误具有普遍性。本部分旨在结合监管规定、执法案例与行业实践，系统梳理这些常见错误，为服务提供者构建合规体系提供经验性指引。

表 18 常见错误与经验总结

错误领域	具体表现	经验总结
对备案/登记义务的判定失误	1. 对服务是否“具有舆论属性或社会动员能力”判断模糊，导致应备案而未备案； 2. 误以为只要调用第三方模型的应用登记即可，但在第三方模型基础上又进行了修改； 3. 误以为“内测”或“公测”并非向公众提供服务，导致应备案而未备案。	1. 并非只有新闻、社交平台才具有该属性，低估了技术的广泛影响力，只要服务涉及向公众提供信息或互动，即可被认定； 2. 仅调用 API（未训练）可申请登记，但如果进行了微调或再训练，则需作为新服务履行备案手续； 3. 现有规定并未豁免测试阶段，只要测试阶段的服务对象是不特定的公众，就会落入监管范围。
数据治理的根本性缺陷	1. 训练语料来源不明，包含未脱敏的个人敏感信息、非法爬取的数据等； 2. 语料库中违法和不良信息比例过高，或缺乏有效的清洗与审核机制。	建立从数据采集、清洗、标注到训练的全流程合规管理体系，为每一类语料留存合法授权证明，对个人信息实施严格脱敏与授权管理。

安全评估报告流于形式	1. 报告内容空洞，未实质覆盖算法透明度、偏见纠正、应急响应等评估要点； 2. 报告内容（如数据来源声明）与实际情况不符，存在逻辑矛盾，构成“虚假陈述”。	安全评估应基于真实情况进行风险排查，建议参照《生成式人工智能服务安全基本要求》，系统性地准备材料，并可考虑引入第三方专业机构协助评估。
没有依法履行备案手续	1. 服务基于未备案的境外开源基础模型开发； 2. 试图通过提交虚假材料包装成合规状态，隐瞒关键技术风险。	应优先选择已备案的基础模型，并要求供应商提供合规证明。任何形式的虚假备案一经查实，不仅面临严厉处罚，更将严重损害企业信誉。
备案材料准备不专业	1. 材料缺失关键文件，或技术文档过于晦涩、未按模板要求填写； 2. 忽视地方监管差异，未按照当地对个别事项的更严格要求调整材料。	1. 对照官方清单逐项核对，并且材料表述应清晰简洁； 2. 主动对接属地网信办，了解并遵循地方具体执行细则。

第五章 附录

一、中央及各部委人工智能重要政策

表 19 中央及各部委人工智能重要政策

序号	政策名称	发布时间	发布主体
1	《国务院关于印发新一代人工智能发展规划的通知》（国发〔2017〕35 号）	2017.7	国务院
2	《教育部关于印发高等学校人工智能创新行动计划的通知》（教技〔2018〕3 号）	2018.4	中华人民共和国教育部
3	《新一代人工智能治理原则——发展负责任的人工智能》	2019.6	国家新一代人工智能治理专业委员会
4	《新一代人工智能伦理规范》	2021.9	国家新一代人工智能治理专业委员会
5	《人工智能安全治理框架》2.0	2025.9	全国网络安全标准化技术委员会
6	《工业和信息化部办公厅关于印发智能制造典型场景参考指引（2025 年版）的通知》（工信厅通装函〔2025〕155 号）	2025.4	中华人民共和国工业和信息化部
7	《国务院关于深入实施“人工智能+”行动的意见》（国发〔2025〕11 号）	2025.8	国务院
8	《国家发展改革委 国家能源局关于推进“人工智能+”能源高质量发展的实施意见》（国能发科技〔2025〕73 号）	2025.9	国家发展改革委、国家能源局
9	《交通运输部 国家发展改革委 工业和信息化部 国家数据局 国家铁路局 中国民用航空局 国家邮政局关于“人工智能+交通运输”的实施意见》（交科技发〔2025〕92 号）	2025.9	交通运输部、国家发展改革委、工业和信息化部、国家数据局国家铁路局、中国民用航空局、国家邮政局
10	《政务领域人工智能大模型部署应用指引》	2025.10	中央网信办、国家发展改革委
11	《关于促进和规范“人工智能+医疗卫生”	2025.10	国家卫生健康委、国家发展改革

序号	政策名称	发布时间	发布主体
	应用发展的实施意见》（国卫办规划发〔2025〕30号）		委、工业和信息化部、国家中医药局、国家疾控局
12	《中共中央关于制定国民经济和社会发展第十五个五年规划的建议》	2025.10	中共中央

二、生成式人工智能相关法律法规

表 20 生成式人工智能相关法律法规

序号	文件名称	生效时间	发布主体
1	《中华人民共和国网络安全法》	2017.6	全国人民代表大会常务委员会
2	《中华人民共和国数据安全法》	2021.9	全国人民代表大会常务委员会
3	《中华人民共和国个人信息保护法》	2021.11	全国人民代表大会常务委员会
4	《中华人民共和国科学技术进步法》	2022.1	全国人民代表大会常务委员会
5	《互联网信息服务算法推荐管理规定》	2022.3	国家互联网信息办公室、中华人民共和国工业和信息化部、中华人民共和国公安部、国家市场监督管理总局
6	《互联网信息服务深度合成管理规定》	2023.1	国家互联网信息办公室、中华人民共和国工业和信息化部、中华人民共和国公安部
7	《生成式人工智能服务管理暂行办法》	2023.8	国家互联网信息办公室、中华人民共和国国家发展和改革委员会、中华人民共和国教育部、中华人民共和国科学技术部、中华人民共和国工业和信息化部、

序号	文件名称	生效时间	发布主体
			中华人民共和国公安部、国家广播电视总局
8	《未成年人网络保护条例》	2024.1	国务院
9	《互联网信息服务管理办法》	2025.1	国务院
10	《网络数据安全管理条例》	2025.1	国务院
11	《人工智能生成合成内容标识办法》	2025.9	国家互联网信息办公室、中华人民共和国工业和信息化部、中华人民共和国公安部、国家广播电视总局

三、人工智能系列相关重要标准

（一）国家标准

表 21 人工智能相关国家标准

序号	标准名称	生效时间	归口部门
1	《人工智能 大模型 第 1 部分：通用要求》(GB/T 45288.1-2025)	2025.2	全国信息技术标准化技术委员会（TC28）归口
2	《人工智能 大模型 第 2 部分：评测指标与方法》（GB/T 45288.2-2025）	2025.2	全国信息技术标准化技术委员会（TC28）归口
3	《人工智能 大模型 第 3 部分：服务能力成熟度评估》（GB/T 45288.3-2025）	2025.1	全国信息技术标准化技术委员会（TC28）归口
4	《网络安全技术 人工智能生成合成内容标识方法》（GB 45438-2025）	2025.9	中央网络安全和信息化委员会办公室
5	《政务大模型应用安全规范》（TC260-004）（技术文件）	2025.9	全国网络安全标准化技术委员会
6	《网络安全技术 生成式人工智能服务安全基本要求》（GB/T45654-2025）	2025.11	全国网络安全标准化技术委员会（TC260）归口
7	《网络安全技术 生成式人工智能预训练和优化训练数据安全规范》（GB/T 45652-2025）	2025.11	全国网络安全标准化技术委员会（TC260）归口

序号	标准名称	生效时间	归口部门
8	《网络安全技术 生成式人工智能数据标注安全规范》（GB/T 45674-2025）	2025.11	全国网络安全标准化技术委员会（TC260）归口

（二）行业标准

表 22 人工智能相关行业标准

序号	标准名称	生效时间	归口部门
1	面向人工智能的数据生产和标注服务能力通用成熟度模型（YD/T 6487-2025）（通信）	2025.11	中国通信标准化协会
2	人工智能开发平台通用能力要求 第 1 部分：功能要求（YD/T 4392.1-2023 ）（通信）	2023.11	中国通信标准化协会
3	人工智能开发平台通用能力要求 第 2 部分：安全要求（YD/T 4392.2-2025）（通信）	2025.12	中国通信标准化协会
4	人工智能开发平台通用能力要求 第 4 部分：大模型开发应用（YD/T 4392.4-2025）（通信）	2025.12	中国通信标准化协会
5	大规模预训练模型技术和应用评估方法 第 1 部分：模型开发 （YD/T 6620.1-2025）（通信）	2025.12	中国通信标准化协会
6	大规模预训练模型技术和应用评估方法 第 2 部分：模型能力 （YD/T 6620.2-2025）（通信）	2025.12	中国通信标准化协会
7	大规模预训练模型技术和应用评估方法 第 3 部分：模型应用 （YD/T 6620.3-2025）（通信）	2025.12	中国通信标准化协会
8	大规模预训练模型技术和应用评估方法 第 4 部分：可信要求 （YD/T 6620.4-2025）（通信）	2025.12	中国通信标准化协会
9	大规模预训练模型技术和应用评估方法	2025.12	中国通信标准化协会

序号	标准名称	生效时间	归口部门
	第 5 部分：模型运营 （YD/T 6620.5-2025）（通信）		

四、生成式人工智能算法备案流程

《互联网信息服务深度合成管理规定》第十九条规定“具有舆论属性或者社会动员能力的深度合成服务提供者，应当按照《互联网信息服务算法推荐管理规定》履行或办理备案、变更、注销等相关手续。深度合成服务技术支持者应当参照前款规定履行备案和变更、注销备案手续。”深度合成服务可参考《互联网信息服务算法备案系统使用手册》和以下指南开展填报工作，具体填报入口和流程如下：

（一）填报入口

登陆互联网信息服务算法备案系统（以下简称备案系统）进行填报，网址为https://beian.cac.gov.cn。系统首页如图 8 所示

图 8 备案系统首页（ 示意图）

（二）填报流程

填报人员需首先注册并登陆备案系统，具体步骤可参考该系统信息公告中的《互联网信息服务算法备案系统使用手册》。登录后的主页面如图 9 所示。

图 9 主页面（示意图）

深度合成备案填报包括三个步骤：一是填报主体信息；二是填报算法信息；三是关联产品及功能信息或填报技术服务方式。

“深度合成服务提供者”（以下简称“服务提供者”）角色 的填报人员需关联产品及功能信息，“深度合成服务技术支持者”（以下简称“服务技术支持者”）角色的填报人员需填报技术服务方式。其中，“服务提供者”是指提供深度合成服务的组织、个人；“服务技术支持者”是指为深度合成服务提供技术支持的 组织、个人。填报流程如图 10。

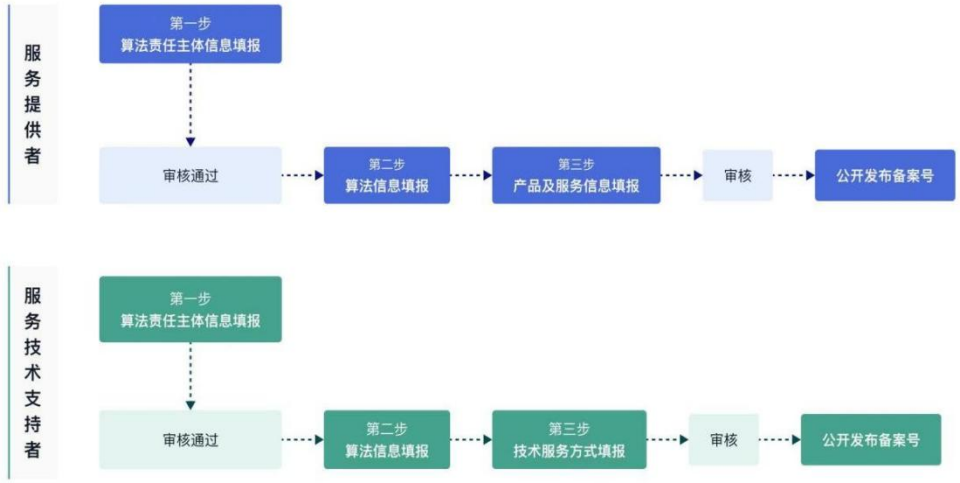


图 10 填报流程（示意图）

1. 主体信息填报

填报人员可点击主页的“主体信息”进行主体信息填报，按照备案主体的实际情况如实填写主体的基本信息、证件信息、法定代表人信息、算法安全责任人信息等内容，并下载附件模板，严格按照模板要求填写并上传《算法备案承诺书》和《落实算法安全主体责任基本情况》附件。主体信息填报页面如图 11 所示。

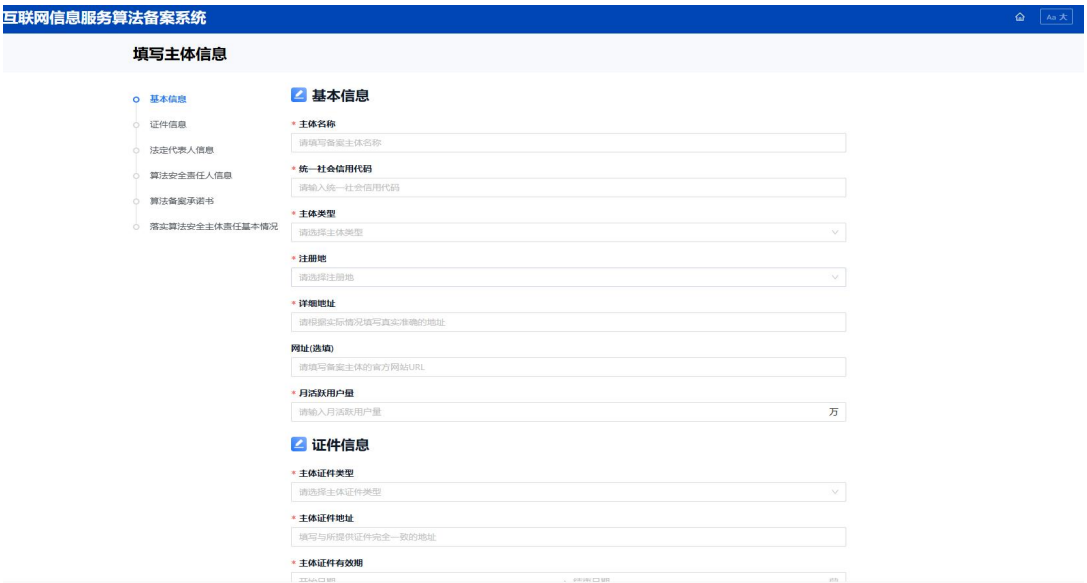
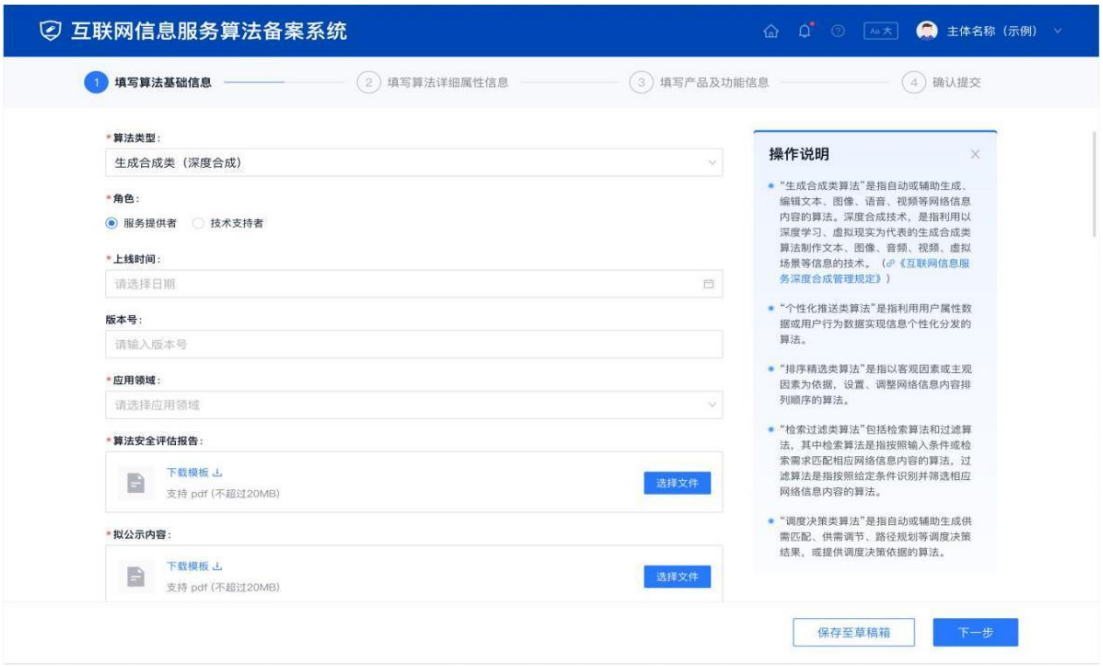


图 11 主体信息填报页面（示意图）

2.算法信息填报

填报人员可点击主页的“备案信息”进行算法信息填报。算法信息填报包括两个步骤：一是填写算法基础属性信息；二是填写算法详细属性信息。

2.1 填写算法基础属性信息



填报人员需选择“生成合成（深度合成）”算法类型，根据实际情况选择“服务提供者”或“服务技术支持者”填报角色。填报人员需下载页面中的模板，按照模板内容填写并上传《算法安全自评估报告》《拟公示内容》等附件。算法基础属性信息填报页面如图12所示。

图 12 算法基础属性信息填报页面（示意图）

2.2 填写算法详细属性信息

填报人员可参考当前填报页面右方的说明文字，根据实际情况填写算法数据、算法模型、算法策略和算法风险与防范机制等信息。填报时，如需中途退出，可点击页面下方的“保存至草稿箱”，保存当前已填写的内容，便于后续继续填写。算法详细属性信息填报页面如图 13 所示。

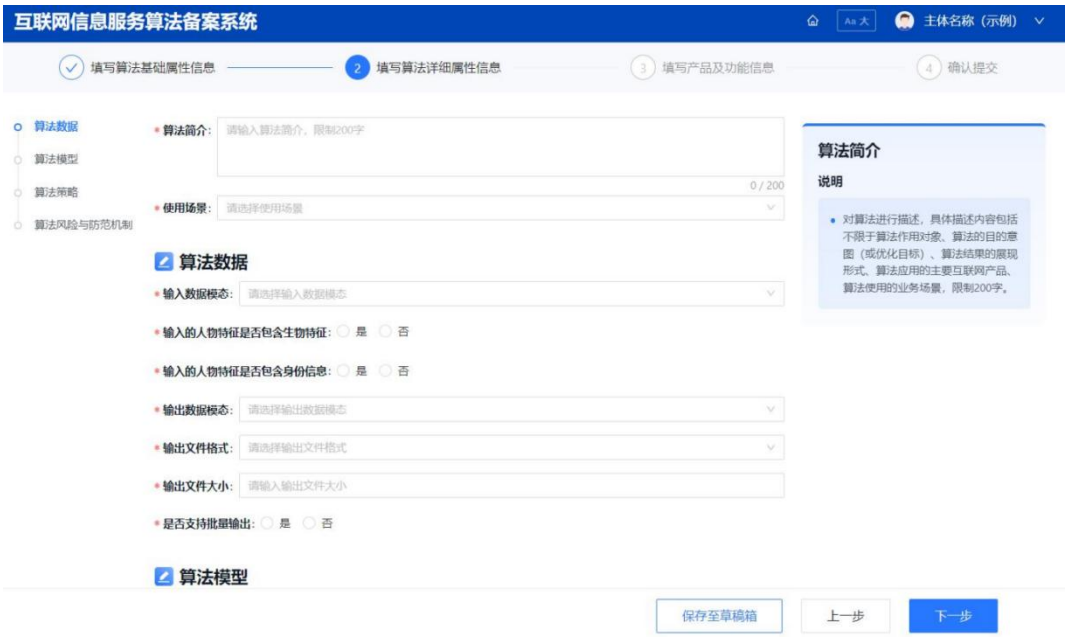


图 13 算法详细属性信息填报页面（示意图）

3.产品及功能信息或技术服务信息填报

在关联产品及功能信息或填报技术服务方式时，“服务提供者”角色的填报人员需关联产品及功能信息，“服务技术支持者”角色的填报人员需填报技术服务方式。

3.1 关联产品及功能信息

“服务提供者”角色的填报人员需根据实际情况勾选应用当前备案算法的产品及功能。需要注意的是，若勾选产品，则表示当前备案算法应用于该产品下所有功能；若勾选功能访问路径，则

表示当前备案算法应用于该路径下所有功能；若勾选特定功能，则表示当前备案算法仅应用于被勾选的功能。勾选产品页面如图14所示。

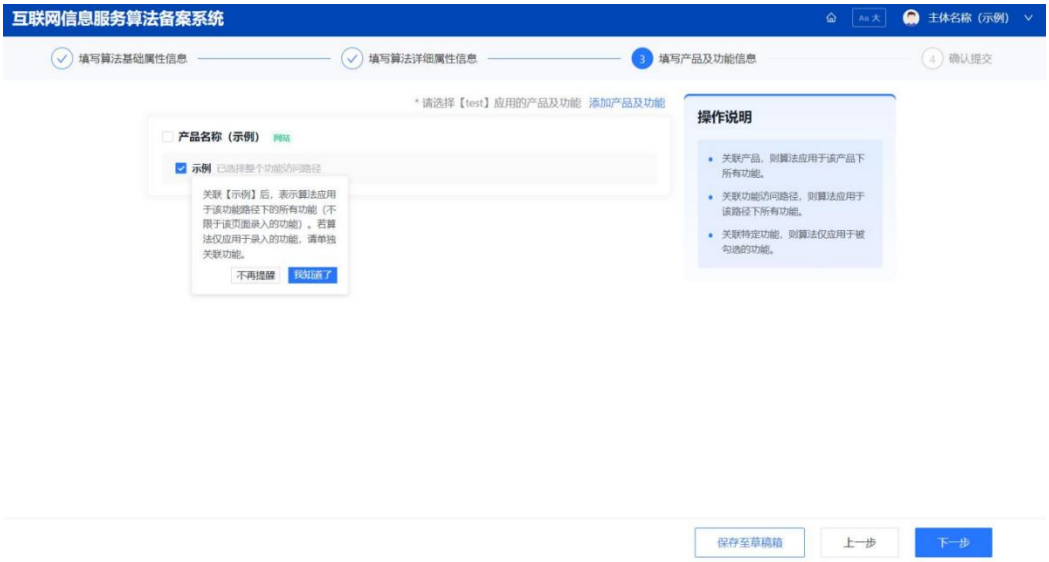


图 14 勾选产品页面（示意图）

若当前产品及功能信息不完善，即产品及功能不能覆盖当前备案算法的关联范围，填报人员可点击该界面下方的“保存至草稿箱”按钮，返回主页并点击主页的“产品及功能信息”完善相应的产品及功能信息。

图 15 产品及功能信息填报页面（示意图）

待产品及功能信息完善后，算法备案填报人员可通过草稿箱返回至“备案信息”页面继续进行算法信息填报。其中，草稿箱的进入方式有两种：一是点击主页右上角用户昵称，在下拉菜单中选择“草稿箱”；二是点击主页“备案信息”按钮，从备案信息界面中进入“草稿箱”。

3.2 填报技术服务方式

“服务技术支持者”角色的填报人员需根据实际情况填写当前备案算法的技术服务方式信息，包括技术服务名称、技术访问方式、技术服务对象、技术服务频度等。填报技术服务方式页面如图 16 所示。

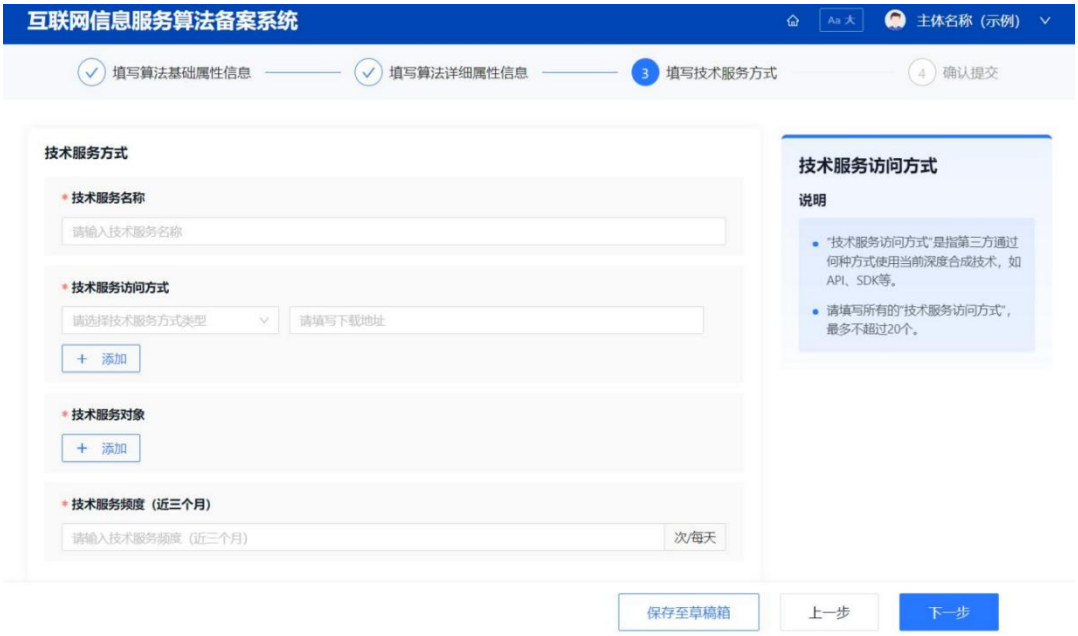


图 16 填报技术服务方式页面（示意图）

4. 提交备案信息

算法备案填报人员在确认填报信息无误后，应勾选页面下方的“我承诺上述填报信息真实有效”，并点击提交按钮，完成算法备案申报。确认信息页面如图 17 所示。

图 17 确认信息页面（示意图）

五、部分省市生成式人工智能服务备案流程（示例）

（一）北京市生成式人工智能服务备案办理流程

第一步：向北京市网信办报备，获取上线采集表

电话联系北京市网信办，说明**生成式人工智能服务**备案意向及模型情况。网信办将通过邮件发送《生成式人工智能（大语言模型）备案信息采集表》。企业填写后连同模型介绍一并通过邮件返回，随后等待审查排队。

第二步：到北京网信办现场测评，获取备案表

北京市网信办对上线备案表和模型情况进行评估，确定需要做**生成式人工智能服务**备案的会联系企业到现场进行测评。测评通过之后，发放《**生成式人工智能服务**上线备案表》等相关文件，该表为核心材料，须认真填写。

第三步：根据表格和评估要点准备材料

在获得备案表之后，企业需要根据备案表的要求和评估要点，系统地准备相应的材料。并组建技术、法务、数据等团队协作准备：

技术团队：①模型描述：提供**生成式人工智能服务**的技术架构、功能特性、算法原理等详细描述。**②服务内容：**明确模型的服务范围、目标用户群体、应用场景等。

法务团队：①服务协议：准备或审查模型服务协议，确保其合法性、公平性和透明度。**②隐私政策：**制定或更新隐私政策，明确用户数据的收集、使用和保护措施。

数据团队：①安全评估报告：准备或更新安全评估报告，包括语

料安全评估、模型安全评估等。**②应急响应预案：**制定详细的应急响应预案，以应对可能的安全事件。

各部门材料需整合审核，确保信息一致、准确、完整。

第四步：向网信办提交材料及测试账号

材料齐全后（包括不限于**备案表、安全评估报告、服务协议、语料标注规则、拦截关键词列表、评估测试题库**等），按要求转化为电子格式，如 PDF 或 Word 文档，连同设置好的测试账号提交至北京市网信办指定渠道。提交后须保持沟通畅通，及时回应审核询问，并根据反馈调整材料或账号设置。及时跟进审核进度，记录并更新相关文档。

第五步：北京市网信办初审

北京市网信办开展形式与实质审查，**如通过：**材料将上报中央网信办复审，企业需与北京市网信办保持沟通，配合补充材料并跟进进度。**如未通过：**根据反馈原因修改材料，如反馈涉及测试账号或模型性能问题，需重新测试后再次提交。

第六步：中央网信办复审

中央网信办对上报的材料进行复审及技术评审，对**生成式人工智能服务**的合规性、安全性等进行全面审查。**如通过：**企业将获得**生成式人工智能服务**备案号，并会进行公示，申请者需在其产品（如网站、APP、小程序等）显著位置标明备案编号并提供公示信息链接，方便查询和监督；**如未通过：**则需重新进行上线备案。

持续合规：持续监控**生成式人工智能服务**的运行，确保其持续符

合备案要求和法律法规。

（二）上海市生成式人工智能服务备案办理流程

第一步：向上海市网信办报备，获取备案表

联系上海市网信办，获取备案材料。可将模型相关介绍以及技术方案发送给上海市网信办审核。

第二步：准备材料并自测自查

依据备案表要求准备材料，**重点自查应答题/拒答题回答率、生成内容合格率、关键词拦截列表、语料来源合法性及安全措施**等环节。

第三步：编写材料并准备测试账号

确保所有材料之间的信息一致。一是检查材料的格式和语言，确保其专业性和可读性。二是准备能够正常使用的产品测试账号。

第四步：向网信办提交材料及测试账号

在完成材料准备和测试账号设置之后，将这些材料正式提交给上海市网信办进行初审。

第五步：上海市网信办初审

上海市网信办初审开展形式与实质审查，可能要求补正材料；材料齐全后，上海市网信办进行线上或线下测试，验证性能与安全性，**如通过：**上报中央网信办复审。**如未通过：**则按意见修改后重新提交。

第六步：中央网信办复审

中央网信办对上报的材料进行复审及技术评审，对**生成式人工智能服务**的合规性、安全性等进行全面审查。**如通过：**企业将获得**生成式人工智能服务**备案号，并会进行公示，申请者需在其产品（如网站、

APP、小程序等）显著位置标明备案编号并提供公示信息链接，方便查询和监督；**如未通过：**则需重新进行上线备案。

（三）广东省生成式人工智能服务备案办理流程

第一步：确定备案级别

根据模型规模、基模属性、服务范围及影响，判断需向国家网信办备案（自研/微调模型）或向省级网信办登记（调用第三方 API）。

第二步：准备申请材料

先进行现场备案预测试，需 2-3 名产品、合规及算法人员参与，建议准备公司及模型介绍 PPT。预测试通过后获取备案材料，主要包括：**生成式人工智能服务上线备案申请表**、**安全自评估报告**、**模型服务协议与隐私协议要求**、**语料标注规则**、**拦截关键词列表**（总规模不少于 10000 个，覆盖 17 类风险，每类风险的关键词不少于 200 个）、**评估测试题集**（覆盖 31 类风险，建议至少每月更新一次）、**测试账号要求**（10 个安全审核账号+10 个无过滤账号）。

第三步：广东省网信办初审

广东省网信办初审开展形式与实质审查，可能要求补正材料；材料齐全后，广东省网信办进行线上或线下测试，验证性能与安全性，**如通过：**上报中央网信办。**如未通过：**则按意见修改后重新提交。

第四步：中央网信办复审

中央网信办对上报的材料进行复审及技术评审，对**生成式人工智能服务**的合规性、安全性等进行全面审查。**如通过：**企业将获得**生成式人工智能服务**备案号，并会进行公示，申请者需在其产品（如网站、

APP、小程序等）显著位置标明备案编号并提供公示信息链接，方便查询和监督；**如未通过：**则需重新进行上线备案。

（四）浙江省生成式人工智能服务备案具体流程

第一步：向浙江省网信办报备，获取备案表

企业需向浙江省网信部门进行备案申请（可电话联系），获取备案表。并准备回答关于模型特点、应用场景及技术优势等初步问询。

第二步：资料撰写、模型内部测试

依据备案表填写材料，详细说明算法原理、数据来源及安全措施。企业内部开展多维度评估与自测，确保符合要求。先将测试账号、电子材料及光盘提交浙江省网信办，根据反馈修改后再正式提交盖章版。

第三步：浙江省网信办初审

浙江省网信办初审开展形式与实质审查，可能要求补正材料；材料齐全后，浙江省网信办进行线上或线下测试，验证性能与安全性，**如通过：**上报中央网信办。**如未通过：**则按意见修改后重新提交。

第四步：中央网信办复审

中央网信办对上报的材料进行复审及技术评审，对**生成式人工智能服务**的合规性、安全性等进行全面审查。**如通过：**企业将获得**生成式人工智能服务**备案号，并会进行公示，申请者需在其产品（如网站、APP、小程序等）显著位置标明备案编号并提供公示信息链接，方便查询和监督；**如未通过：**则需重新进行上线备案。

（五）江苏省生成式人工智能服务备案办理流程

第一步：向江苏省网信办报备，获取备案表

企业需向江苏省网信部门进行备案申请（可电话联系），获取备案表。按照备案要求，准备所有材料，包括**生成式人工智能服务上线备案表、安全评估报告、模型服务协议、语料标注规则、拦截关键词列表、评估测试题集等**。

第二步：提交备案申请

填写备案表并提交全套材料至江苏省网信办。

第三步：网信办初审

江苏省网信办初审开展形式与实质审查，可能要求补正材料；材料齐全后，江苏网信办进行线上或线下测试，验证性能与安全性，**如通过：**上报中央网信办。**如未通过：**则按意见修改后重新提交。

第四步：中央网信办复审

中央网信办对上报的材料进行复审及技术评审，对**生成式人工智能服务**的合规性、安全性等进行全面审查。**如通过：**企业将获得**生成式人工智能服务**备案号，并会进行公示，申请者需在其产品（如网站、APP、小程序等）显著位置标明备案编号并提供公示信息链接，方便查询和监督；**如未通过：**则需重新进行上线备案。