

# 黄殷颖： 永远有目标，永远在路上



□本报记者 张仁平  
通讯员 雷阔娟 薛文洪

“点击福州数字检察展示平台上的相关业务模块，各类数据可自动生成并呈现在眼前；输入关键词，监督线索即能从海量的数据池中‘跃然而出’。”今年1月16日，福建省福州市检察院检察技术信息部主任、数字办负责人黄殷颖正演示着一个科技范儿十足的法律监督模型。作为同事眼中勇往直前的“闯将”，近年来，他带领本院数字检察团队一步步攻坚克难，用数字检察赋能法律监督提质增效，助力优化“三个管理”，为福州市检察机关数字检察工作写下浓墨重彩的注脚。

## “两把刷子”敲开“数字之门”

“本科计算机专业毕业，研究生攻读法学”，在该院领导和同事看来，黄殷颖是开拓数字检察工作的“不二人选”。2022年，他离开工作了十余年的办公室岗位，转任该院检察技术信息部主任兼数字办负责人。

面对新技术门槛和新领域的挑战，如何担起全市检察机关数字检察起步开局的重任？黄殷颖迅速调整角色，迎难而上，从学习数据分析到钻研复杂算法、从演练数据抓取到进行检察业务实操，他始终保持强烈的求知欲。

为了提升专业能力，他多次赴省外先进检察院学习取经，与福建省内高校大数据研究所、相关科技企业交流合作，并深入基层检察院和相关企业，与基层一线检察人员交流监督办案逻辑方向和数字化需求。学习的场景不断切换，也促使他快速掌握了数字检察的核心要旨，拓宽出将数字技术融入检察业务的新视角。

本领练就了，如何进一步实现大数据与检察工作的深度融合？面对新问题，他又主动当起数字检察理念的传播者，在全省检察技术信息工作培训班上向全省参训人员授课，在福州市检察机关，更是通过数十次的专

题讲座、数字检察工作例会、座谈会、业务交流会等，用深入浅出的话语和具体案例讲解，消除大家对数字浪潮的茫然心态。

在该院党组推动和黄殷颖的带动下，数字办案思维和理念逐渐在福州市检察机关生根发芽。基层一线办案人员成为数字检察主力军的态势逐渐成形，许多办案检察官开始从典型案例甚至日常普通案件中分析、梳理、提炼类案规律及背后存在的社会治理漏洞，提出数字监督思路和数字模型建设方向，数字赋能办案逐渐成为一种履职自觉。“于我而言，看到办案人员主动将数字思维运用到办案实践，建立起简单的数据模型有效应用于办案时，收获感无以言表。”黄殷颖说。

传帮带的实践并未止步于省内外。近两年来，他先后应邀到省外检察院系统授课10余场，全面分享数字检察的理念思路、路径方法、实务操作和应用成果，并赴新疆维吾尔自治区昌吉回族自治州检察院开展对口数字检察援助工作。

## 让数据成了办案“好帮手”

“为什么重大涉税违法失信案件公开信息中，移送司法机关的信息较多，而我们检察业务应用系统中可查询到的涉税案件却较少？”黄殷颖在查询政府公开网站信息时，发现本地税务机关近5年公示的重大涉税违法行政处罚近千件，数据分析显示，标注移送司法机关的有300余件，但通过查询检察业务应用系统发现，全市近5年仅有6件涉税犯罪的起诉案件，成案率不足2%。

“消失”的案件去了哪里？黄殷颖和同事分析发现，问题很可能出在税务机关案件材料移送、公安机关立案等环节，“这个数据差也指出一条监督线索方向。”但数据上的疑点还需办案检察官进一步调查核实，才能反映出真实情况。为此，他带领技术团队与该院经济犯罪检察部的办案团队密切协

作，不断完善数据分析逻辑，认真研究涉税犯罪行为要素，构建起“行刑衔接”涉税领域立案监督大数据模型，重点关注税务机关案件材料移送、公安机关立案等环节，让检察官通过一键筛选便能从海量的数据中发现监督线索。

在分析线索时，一起福州市某公司虚开增值税专用发票涉税金额巨大的信息进入黄殷颖的视野，“案发地区属于进出口贸易发达地区，是否还存在类似企业？”调整数据分析逻辑后，他运用该模型批量发现该地区存在的7件类似情况。该院业务部门将线索移送基层院调查核实，最终立案监督7件，推动追赃挽损3600余万元。因模型取得较好应用成效，被省内外10多个地市检察机关借鉴应用。

“我们想要的监督点，他都能以数字化的形式帮我们实现，有效提升了办案效率。”数字赋能法律监督和检察管理带来的质效，该院经济犯罪检察部主任张昭兰感受尤深，在应用模型过程中，福州市检察院业务部门对数据统一汇聚、归集、分析，将监督线索统一分解下发各基层院办案一线，并实时跟踪、指导、督促，在减轻基层院数据收集整理压力的同时，也凝聚起上下一体的履职合力。

## 逐浪“数字蓝海”勇立潮头

让更多“沉睡数据”被充分激活和运用，是黄殷颖一直努力的方向。

“职务侵占案件已经判决生效了，为什么我们查不到返还公司损失的判决执行情况？”2024年4月，该院在走访企业时，黄殷颖听到企业负责人的愁盼。

2022年8月，这家贸易公司业务夏某利用职务便利，将公司100余万元的货物低价卖给他人。该案除刑事

判决外还涉及企业损失赔偿款31万元，但该款一直未执行。

如果生效的判决未依法执行，判决书对涉案企业来说就是一张“空头支票”。黄殷颖针对侵害企业合法权益犯罪案件刑事裁判生效后未及移送执行、财物处置进程缓慢等问题，对2020年以来涉及合同纠纷、职务侵占、挪用资金等案件，建立数字模型进行全面筛查，形成了涉企追赃挽损专项数据池，并通过对比执行信息碰撞比对，发现255条案件线索。

据此，福州市检察机关在2024年第二季度部署开展数字检察小专项行动，逐一列明台账，逐项排查核实，发出检察意见67份，其中督促执行立案55件，涉及金额1800余万元，督促恢复执行24件，涉及金额1600余万元，有力维护企业合法权益。

“数字赋能破解涉企追赃挽损执行难问题，给了我很大启发。数字检察要深入推进，就要融入党委政府中心工作，全市检察机关共同发力开展专项监督，才能有更大的规模效应、更好的治理效果。”黄殷颖说。2024年，他带领数字检察团队，聚焦检察机关主责主业和上级检察机关的部署要求，围绕企业、民生、安全等主题，采取“数据统采分用”“台账制管理”等新路径，部署开展了数字检察9个专项监督。

“新的一年，我们将加强数据管理能力，探索构建反走私行刑共治平台、业务态势感知体系和相关罪名案件自动筛查系统，不断挖掘大数据‘红利’，赋能检察监督办案水平。”黄殷颖瞄准新的目标，开启了新的征程。

但对合成数据的使用，必须保持谨慎态度，尽量减少错误，确保其作为真实数据的可靠补充，从而保障人工智能模型系统的准确性和可信度。

——摘编于《科技日报》



大语言模型，是基于深度学习技术构建的，能够处理和生成自然语言文本的复杂系统。这类模型的运行原理是通过在大量文本数据上进行训练，学习语言的模式和结构，从而实现对话的理解和生成。它们通常包含数亿甚至更多参数，能够捕捉语言的细微差别，并在多种任务中表现出色，如文本分类、机器翻译、问答系统等。

大语言模型的核心优势，在于其强大的上下文理解能力，能够根据对话或文本的上下文提供准确的回应，从而生成连贯、逻辑性强的文本。但大语言模型也面临挑战，包括对训练数据的依赖、潜在的偏见问题以及弱解释性。为了提高模型的性能和可靠性，还需要采用高质量的训练数据、透明的算法设计和持续的模型评估。

近年来，AI技术实现了飞跃式进步。在一定程度上，AI技术像一个虽还在学习走路，但已成为思维敏捷、知识渊博、理解力强的助手。特别是以DeepSeek为代表的大型语言模型，其发展速度迅猛，呈现出AI技术在完成自然语言理解、文本生成、数据分析等复杂任务时具备的显著能力。

伴随着技术的发展，检察机关也在探索利用数字技术对检察工作进行全面和系统的改革与升级。面对AI技术的快速发展，数字检察既遭遇了新的挑战，也迎来了新的发展机遇。问题的关键在于，数字检察如何把握科技机遇，形成新的发展格局。

长期以来，数字检察在应用AI技术时，主要关注的是如何利用这些技术处理和分析海量的检察数据。部分典型应用包括利用深度学习等自然语言理解技术，自动地从大量案件文本中抽取争议焦点、法律适用等关键要素，实现更加精准的类案推荐；通过对特定案由案件的大样本深度分析，构建类案知识图谱，提出更加规范恰当的量刑建议；对虚假诉讼案件进行实证分析，包括当事人、代理人、行为模式、事实、背景等特征，以此构建虚假诉讼预警模型，自动筛查海量已决案件裁判文书中的虚假诉讼风险等。这些应用已取得一定成效，但从技术角度看，AI技术的潜力并未得到充分挖掘。目前的应用模式本质上还是以检察人员的个人经验和认知为中心，AI技术更多是作为辅助完成简单重复性工作的工具。尽管这种应用大幅度节省了人力成本并提升工作效率，但从本质上来看，AI技术的作用尚未达到不可替代的程度。

大语言模型的引入使得AI技术的效能数字检察中得到进一步释放。相较于传统AI技术，大语言模型在上下文理解、自然语言生成、多功能性、交互体验、可扩展性等诸多方面具有显著优势。只要具备足够的高质量语料用以模型训练和调优，大语言模型不仅能够更好地完成类案推荐、量刑建议、虚假诉讼风险识别等传统任务，还能协助检察官高效处理案件文本，审查关键信息，并在检察工作的各个环节提供决策建议和相关信息，同时自动生成各类文书。因此，大语言模型与以往的AI应用有所不同，它有望成为检察官真正的随身助手，全面赋能检察工作各环节。

此外，大语言模型的创新应用还能进一步拓展数字检察的辐射范围。当前，社会科学研究者已经开始利用AI模拟平台在虚拟环境中设计和仿真社会实验，通过虚拟AI代替真人进行研究，此方式可避免真实社会实验所需的高昂成本和引发的社会伦理问题。在检察领域，大语言模型同样可以用于多种场景的模拟和仿真。例如可模拟真实的办案场景，为新入职的检察官提供更为真实的培训体验；模拟新的检察制度和法规政策的实施场景，研判施策方案和效果，从而提高相关规定的前瞻性和有效性。

目前，针对检察机关的大语言模型研发和应用尚处于初级阶段，其输出内容尚未满足检察办案的核心需求。进一步发挥大语言模型在数字检察中的作用，可从以下三个方面着手：一是建设高质量语料库。当前与检察相关的高质量语料较为稀缺，这限制了模型理解和生成法律领域文本的能力。检察机关可以通过建设专门的检察语料库来解决这个问题，不断收集和整合包括法律法规、司法解释、法律文书、法学研究文献、法律新闻报道和评论、各类媒体平台信息等多种类型的语料。二是完善数据开发利用机制。模型训练需要大量语料数据，其中的原创成果和个人信息可能引发知识产权保护、个人信息权益保障等方面的风险和争议。当前，对数据确权等基础问题的理解还存在争议，但从实际需求出发，从政策层面突破制度障碍，先行先试，以促进高质量语料数据的开发和利用。三是应对司法伦理问题。在应用大语言模型过程中，应当始终坚持司法为民的基本立场，确保AI的应用不偏离实体公正和程序公正的基本要求，努力让人民群众切实感受到司法公平正义。

（作者为东南大学法学院副教授、博士生导师）

# 与大语言模型“交手”

□冯煜清

# 训练模型的“养分”从何而来

以往，人工智能模型主要依赖真实的数据来构建、训练和改进，而真实数据多由公众创建的文本、视频和图像构成，科技人员通过调查、实验、观察或挖掘网站及社交媒体等途径收集。随着人工智能的快速发展，公众生成数据的速度已无法应对人工智能模型不断增长的需求，真实数据量略显匮乏。

为给人工智能模型构建提供充足的“养分”，科技人员正转向使用由算法

生成、可以模仿真实世界的合成数据。合成数据在数字世界中创造生成，作为真实数据的替代品，可用来训练、测试、验证人工智能模型。合成数据可无限供应，为训练人工智能模型提供经济高效且快捷的解决方案，同时有效解决了人工智能训练使用真实数据时面临的个人隐私和敏感信息保护等问题。

虽可解燃眉之急，但也并非尽善尽美。当人工智能模型过于依赖合成

数据时，它们会产生更多“幻觉”，编造看似合理可信但实际上并不存在的信息，导致模型的质量和性能下降。此外，合成数据也存在过于简单化的风险，它可能缺乏真实数据集蕴含的细节和多样性，这可能导致其训练的人工智能模型过于简单，缺乏充足的实用性。人工智能模型的发展在很大程度上取决于数据质量，虽然合成数据在克服数据短缺方面发挥重要作用，

但对合成数据的使用，必须保持谨慎态度，尽量减少错误，确保其作为真实数据的可靠补充，从而保障人工智能模型系统的准确性和可信度。

——摘编于《科技日报》



# 让无证驾驶“驶”不得

上海市奉贤区检察院检察官在办案中发现，被吊销驾照的社区矫正对象违反规定无证驾驶的现象并非个例。为强化源头治理，该院构建社区矫正对象无证驾驶情况刑执监督模型，以辖区内危险驾驶罪或交通肇事罪、依法被吊销驾驶证的社区矫正对象数据信息为基础，精准查找、及时发现社区矫正对象无证驾驶违法行为。截至目前，该院利用模型排查发现9人存在无证驾驶情况，其中4人因情节严重被撤销缓刑，收监执行。同时，该院与相关部门共建专项工作机制，推动开展常态化整治工作。图为该院检察官到上海市公安局奉贤分局交警支队开展数据比对和线索核查。

（本报记者陈洁婷 通讯员孙晓光/图文）

# 数智化推动网络餐饮食品安全数治化

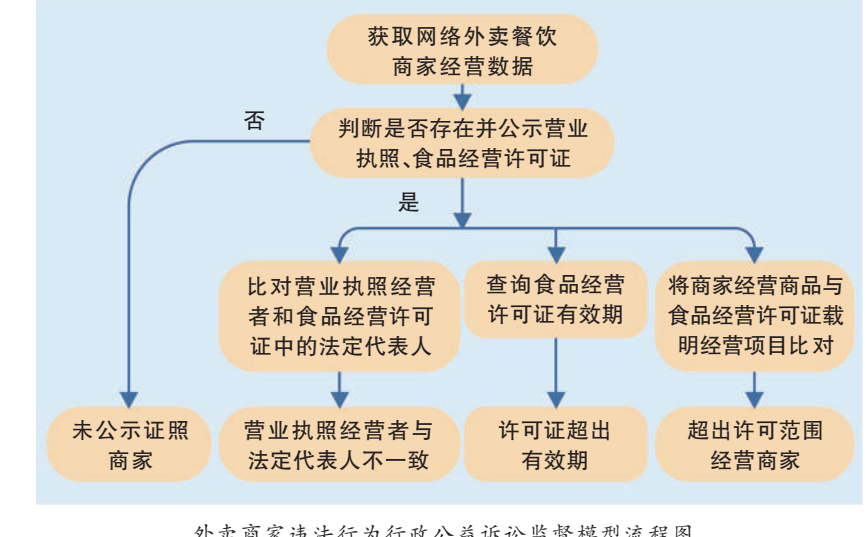
□谢飞 韩恒

近年来，网络餐饮行业发展迅速，但也存在经营不规范等现实问题。安徽省淮北市检察院在贯彻落实最高检部署的“保障千家万户舌尖上的安全”专项监督活动中，发现网络餐饮外卖平台对淮北市辖区入驻商家监管不严，导致部分资质不全、存在食品卫生隐患的商家在平台经营，危害食品安全，相关行政管理部门可能存在履职不到位的问题。

为此，淮北市检察院将“网络餐饮服务生产者食品加工违法行为及监管部门是否依法履行监管职责”作为监督工作的重点，积极践行数字检察理念，依托全国检察机关大数据法律监督模型管理平台，构建外卖商家违规行为行政公益诉讼监督模型。该院公益诉讼检察部门依据模型筛查线索并开展调查核实后，向市场监督管理部门制发检察建议，取得较好的社会治理效果。

## 【模型构建】

数据来源：网络外卖平台餐饮商家经营数据，包括商家店铺经营信息、营业执照信息、食品经营许可证信息。其中，从商家经营页面获取餐饮商家经营名称、经营范围、地址等信息；从商家食品安全档案公示页面获取证照信息，包括营业执照名称、经营者、经营范围等信息；从商家食品经营许可证中获取经营者名称、法定代表人、经营项目、证件有效期等信息。数据获取后，将上述数据要素分别存储到对应的数据表中。



外卖商家违法行为行政公益诉讼监督模型流程图。

## 【数据分析】

第一步：比对商家营业执照和食品经营许可证信息，判断商家是否上传相关证件，输出未公示证照商家线索。

第二步：比对同一商家营业执照经营者与食品经营许可证中载明的法定代表人是否一致，输出证照经营者与法定代表人不一致的商家线索。

第三步：将商家食品经营许可证有效期与当前日期进行比对，输出食品经营许可证超出有效期的商家线索。

第四步：将商家经营商品与食品经营许可证载明的经营项目比对，输出超出许可范围经营商家线索。

第五步：对上述输出的线索进行

核查，如果满足其中之一，即视为外卖商家存在违规行为，存在食品安全隐患。

## 【应用成效】

淮北市检察院应用该模型，发现全市范围内网络外卖平台餐饮商家在食品经营活动中的违法违规线索400余条，共成案19件，有效助力公益诉讼检察部门督促相关行政机关依法履责，保障人民群众舌尖上的安全，促进了行业健康发展。具体成效体现在以下四方面。

一是全面筛查数据，发现违规商家。该模型在上线运行后，该院对淮北市辖区的1500余家网络餐饮商家信息筛查后，发现30余个存在违规行为的商家，其中多个本地知名商家在不

同程度上存在违规行为。该院公益诉讼检察部门依此线索开展工作，先后向市场监督管理部门制发5份检察建议，全部得到采纳。

二是精炼建模思路，提升监督效能。该院在模型构建中，坚持“业务+技术”双向发力，前期多次碰撞研究，梳理简化模型构建核心要素，坚持“少而准”原则创建建设路径。模型通过抓取、分析数据，快速筛选出存在违规行为的商家，使得检察人员能够有的放矢地进行核查和处理，避免了传统监管方式中过多的人力消耗和精准性不足的问题。通过模型的应用，推动实现了对网络餐饮商家的全方位、全时段监督，不仅提升了监管质效，也节约了司法资源。

三是建立协作机制，共筑食品安全防线。该院以此为契机，与市场监督管理部门建立监督通报、联席会议、线索移送、联合调查等协作机制，推动辖区食品安全领域健康、安全、高质量发展，回应人民群众对食品安全的新期待。

四是推广应用模型，推动行业治理。因模型应用成效显著，安徽省在全省范围内以专项行动的形式对该模型应用于推广，取得良好成效。该模型在上架全国检察机关大数据法律监督模型管理平台后，全国数百家检察院累计下载应用900余院次，有力推动了网络餐饮外卖食品安全行业治理。今年1月，该院因积极推广该模型，获评全国检察机关大数据法律监督模型推广活动“模型推广优胜院”。（作者单位：安徽省淮北市检察院）