

按照最高检关于省级院要稳步建设大数据法律监督模型管理平台的要求,安徽省检察院肩负起省级检察机关对数字检察工作的规划和统筹职责,切实将数字检察法律监督平台和应用的建设的重点落脚于业务办案实践——

# 以效果为导向构建集约化数字监督平台

□鲁建武 樊静 张大伟

检察机关推进数字检察战略是以检察工作现代化服务中国式现代化的重要举措。为落实最高检加快推进数字检察工作部署,安徽省检察院按照“党组抓总、业务主导、技术跟进、典型示范、争创一流”的工作思路,以“通过大数据应用于办案实践,发现一类问题,办成一类案件,促进一域治理,形成数字赋能检察监督的安徽经验”为目标,积极探索开展数字检察工作,现已建成涵盖六大中心的数字检察法律监督平台及应用。

### 平台建设初衷

数字检察战略是法律监督手段的“革命”,是提高法律监督能力的重要依托。当前,安徽省检察机关在数字检察思维的引导、工作机制的建立、数据治理和归集等方面积极开展工作,鼓励基层检察院先行先试,“小、精、专”的应用场景不断涌现,在模型创新和应用上取得初步成效。但由于缺少可供实现数据治理、数字检察办案省级使用的平台,致使市、县(区)检察院模型创新起步难、数据共享难、经验推广难,容易造成低水平重复建设、形成信息孤岛,不能及时形成“一域突破、全省推进”的示范效应。为此,根据最高检关于省级院要稳步建设大数据法律监督模型管理平台的要求,亟须建设数字检察法律监督平台。

### 平台建设思路

安徽省检察院认真落实最高检关于“业务主导、数据整合、技术支撑、重在应用”的数字检察工作机制要求,肩负起省级检察机关对数字检察工作的规划和统筹职责,切实将数字检察法律监督平台和应用的建设的重点落脚于业务办案实践。

建立全省三级检察机关统一的自主创新开放式平台。全省统建平台能够实现项目集约化,统一进行管理、运维,

避免各地市各自为政、浪费资源,能够为全省各级院提供数据接入、存储、治理服务,帮助开展全省各级检察院监督线索研判,辅助全省各级检察院进行监督模型的创新。平台立足于解决在数字检察工作中数据没地方放、没办法开展分析、创新监督思路没法实现等问题,避免全省三级检察院低水平重复化建设、重复化运维。

反复对接基层、对接业务部门,在平台上建立初始模型和应用。安徽省检察院在反复对接基层检察院和相关业务部门的基础上,整合吸收各类关于业务监督点、监督规则、试用体验、交互界面等方面的意见建议,稳步开展安徽数字检察法律监督平台建设。初始阶段,在平台建立涵盖“四大检察”的10个模型和应用,这样既能满足全省短期内开展数字检察工作需要,又能针对今后市县检察院新增的应用进行灵活扩充、敏捷开发,更好地服务基层检察机关办案。同时,平台还预留升级接口,能够根据数字检察工作开展情况,不断迭代升级。平台采用模块化开发方式,设计通用接口,方便接入优秀法律监督模型,能对相关功能进行升级和拓展,不断满足全省三级检察院对法律监督工作的需求。

强化数字检察法律监督平台安全性建设。数字检察法律监督平台包含了检察机关办案数据、系统研判的办案线索、行政机关及司法机关的数据等诸多信息内容。平台在网络安全、数据安全层面进行重点建设,做好系统安全性、数据安全、身份认证及访问控制、系统备份及恢复等层面的加固和保障,提升针对非法入侵、数据泄露、恶意攻击等风险隐患的防御能力。平台建设了入侵检测、数据库防火墙、数据库审计等安全设施,不断加强对相关人员的安全培训,确保数字检察法律监督平台应用运行稳定。

### 平台建设内容

安徽省检察院坚持问题导向、效果导向,在开展数字检察工作中,因地制宜,

边摸索边推进,现已建设涵盖六大中心的数字检察法律监督平台。

薄平台、重应用,建成涵盖六大中心的法律监督平台。安徽省检察院采用“平台+应用”的建设思路,构建数字检察“1+N”模式。“1”是按照“薄平台、轻建设、重应用、重办案”的理念,先统建满足年度目标任务的基础性数据平台,将平台全部的建设任务按需分解至未来几年,将数字检察工作的重点放在法律监督模型的创新和平台线索的发现上,利用平台促进线索成案,深入推进数字检察赋能案件的办理。“N”是指在此平台上建设包括刑事立案监督等多个大数据法律监督模型和应用。

其中,六大中心包括数据治理中心、线索研判中心、办案指挥中心、模型创新中心、学习研究中心和能力支撑中心。数据治理中心主要负责多渠道获取的海量数据进行有效管理和监督,以确保数据合规和安全;线索研判中心则是利用人工智能技术对海量数据进行分析和研判,快速发现线索、确定关系,辅助检察官提高研判效率和准确度;办案指挥中心旨在通过对平台中线索的指挥和调度,帮助数字检察专班人员掌握线索情况,指导办案,确保线索流转的高效有序;模型创新中心通过轻松简便的可视化界面,让检察官更加专注于模型创意而不是复杂代码。在比较准确的监督规则和相对量的数据基础上,对于类似对比查找、分类汇总、筛选提取等简单法律监督模型,只需鼠标进行拖、拉、拽就能快速构建,对于较为复杂的法律监督模型则可以通过多个监督模型进行调用组合来逐步实现;学习研究中心聚合了数字检察相关视频教程、学习资料、期刊论文及互动论坛等,主要供全省检察人员学习研究、交流经验;能力支撑中心是实现上述五大中心功能所需要具备的底层支撑,如OCR识别、自然语言处理、自动化推理和推断等。

“自我原创+外部引进”,甄选数字检察监督模型应用。数字检察监督模型,是数字检察法律监督平台的重要组成部分。经多次调研,安徽省检察院既

积极吸收外省较为成熟的模型应用,如刑事案件下行监督、民事生效裁判监督等,又结合安徽省开展的刑事立案监督、行政检察“空壳公司”治理等相关专项工作,并充分考量该省办案检察官对数字检察监督办案创新的接受程度,最终梳理出涵盖“四大检察”的各类检察业务监督模型20个。安徽省检察院对这些模型的具体监督规则和数据分析规则进行深入研究,召开大数据法律监督视频研判会,在省、市、县三级检察院的共同推动下,结合全省检察工作实际,在数字检察法律监督平台中研发建设监督模型10个,集中在刑事审判监督、刑事执行监督、民事检察、公益诉讼检察、未成年人检察等业务条线。

强化管理,注重保护,积极做好平台的数据治理。数据是数字检察监督平台的关键脉络,是发掘线索的重要基石。平台将各个渠道获取的不同数据,依据数据类别采取不同的治理流程及措施,根据统一制定的数据质量管控标准、数据保密级管控标准,对数据进行分类管理。数字检察法律监督平台制定了完善的数据标准管理、数据质量管理、数据保密管理等策略,保障数据的可靠性、可维护性和安全性。在检察机关内部,平台根据数据特征及各数据应用系统的数据要素需求,在原始数据库之上维护融合库与主题库,用于数据快速检索、碰撞、维护。对平台外接的监督应用,平台统一以接口形式提供基于安全策略管理的数据服务,以满足各市级检察院、基层检察院的个性化法律监督需求。

目前,安徽省检察院建成的数字检察法律监督平台已大力支撑全省大数据法律监督专项行动,平台已就契合此专项行动10个专项监督进行了应用构建。检察官对平台推送的线索进行研判分析,办结相应案件,实现了发现一类问题,办成一类案件,促进一域治理。

(作者分别为安徽省人民检察院副检察长、数字检察办公室主任,安徽省人民检察院检察信息技术部主任、数字检察办公室副主任,安徽省人民检察院检察信息技术部主任科员)



深圳市光明区检察院开展“虎牙”系列行动,对酒店行业工作人员开展普法宣传。

□樊江 华夏

近年来,全国各地性侵害未成年人刑事案件总体呈上升趋势。广东省深圳市检察机关未检工作从精准帮教、精准保护迈进精准预防的新阶段,在开展综合履职过程中,发现有部分未成年人被侵害的场所发生在酒店。为支撑监督履职,深圳市检察院与深圳光明区检察院联合开发酒店行业未保线索大数据法律监督模型,有效推动深圳市检察机关全面铺开“虎牙”系列未成年人保护法律监督专项工作。“虎牙”系列聚焦酒店、酒吧、KTV等领域,以“四大检察”综合履职为抓手,以预防和减少未成年人被侵害为目标,通过大数据赋能开展“四检合一”工作,解决了以往监督履职碎片化、监督效能短期化等问题,实现了在业务主导下,有效整合监督数据,形成可持续的监督流,取得明显监督效果,实现由“治罪”向“治理”再向“治理体系”的发展,促进了酒店行业良性发展。

### 【模型构建】

数据来源:构建酒店行业未保线索大数据法律监督模型所使用的数据,不仅包括检察业务应用系统中刑事案件涉及的酒店、宾馆的数据,还包括深圳智慧未检系统中司法社工跟进精准帮教和精准保护案件的数据、深圳市未成年人入住酒店旅馆的数据以及公安机关对相关酒店的行政处罚数据等。

数据清洗:按照案发地(酒店)、案由(强奸等)、被侵害对象(未成年人)三个维度,采取关键词抓取、数据碰撞、权重赋分(即把不同的关键词根据重要程度标识不同分值,分值越高表示越重要,以此识别出有效数据)等技术,依托数据清洗的技术方式,筛选挖掘出各类监督线索,为开展未成年人保护综合履职监督奠定基础。

数据模型:数据模型通过两组不同数据之间的比对,以及对初步比对的数据再次与第三组进行比对的方式,找到各类监督价值且可以开展监督履职的数据。一是通过对刑事案件数据进行初步筛选,将涉案酒店与公安机关系统中的入住酒店记录进行比对、碰撞,将涉案酒店违法接待未成年人的违法行为相关数据与公安机关的行政处罚数据进行比对,判断是否存在未依法处罚涉案酒店的情形,从而推送行政公益诉讼线索;二是将涉案酒店违法接待未成年人的相关数据与公安机关的行政处罚数据进行比对,判断是否存在因执法机制缺陷致使侵害公共利益的情形,从而推送未检行政公益诉讼线索;三是通过筛选出的存在违法接待未成年人的涉案酒店数据与酒店工商登记数据进行比对,发现涉案酒店侵害未成年人公共利益的程度,判断是否存在侵害不特定未成年人公共利益的情形,从而推送民事公益诉讼的线索;四是根据将在酒店被性侵害的未成年人及其监护人的数据与法院民事侵权受案立案的原告信息进行比对,从中发现是否存在被侵害未成年人的监护人未提起民事侵权诉讼的情形,进而推送未检民事支持起诉线索。

### 【应用实效】

酒店行业未保线索大数据法律监督模型在深圳市检察机关开展的“虎牙”系列未成年人保护专项工作中发挥了重大作用,取得显著成效。

一是实现检察能动履职,监督方式由单一履职变成“四检合一”综合履职。深圳市检察机关不断深化未成年人刑事、民事、行政、公益诉讼“四大检察”融合履职,综合运用“四大检察”职能,实现未成年人检察业务集中统一办理。以酒店住宿业违规接纳未成年人入住作为切入点,将一体化办案模式融入家庭保护、学校保护、社会保护、网络保护和政府保护中,提高监督效能,共完成各类监督立案165件,2022年下半年全市发生在酒店的侵害未成年人刑事案件数量同比减少51.5%,其中光明区减少85%。“四检合一”的优势真正实现从个案到类案,再到社会治理的检察监督模式转变。

二是依托大数据赋能,由传统治罪向数字治理转变。传统治罪模式主要是依法追究犯罪人的刑事责任,而对被侵害的未成年人开展的保护工作,属于治已病的末端保护,弊端在于保护难度大且成本高。而数字治理依托大数据将未成年人被侵害的原因进行信息化解析,便捷化还原呈现了未成年人被侵害的精准原因,为检察监督履职指明了明确方向,提高了监督的力度和效果,避免更多未成年人被侵害,属于治未病的前端保护,达到“最好的保护就是预防”的目的,以大数据为驱动力,最大程度激发数据对未成年人检察监督的叠加、倍增作用,赋予检察监督依法能动履职的全新方式和手段。检察官不再囿于就案办案,而是不断开拓思路,实现了“四大检察”综合履职。

三是拓展行业治理模式,构建数字治理体系。深圳市检察机关立足检察职能推动实现“预防优先”的行业治理范式。在开发适用酒店行业未保线索大数据法律监督模型的基础上,充分利用大数据思维,将监督模型拓展到向未成年人售烟售酒、违法给未成年人纹身、不履行强制报告义务等未成年人保护等专项行动中,积极与市场监督管理局、市文化广电旅游体育局、市卫生健康委员会等相关部门探索合力共建数据治理体系,实现数字化时代多部门共建共治共享的综合治理大格局,最大程度实现对未成年人的综合保护。

(作者单位:广东省深圳市光明区人民检察院)

## 近看模型

# 找准路径让业务数据“活”起来“稳”下来

## 业务数据资源化

□赵宝柱 张爱丽

数字时代背景下,如何有效运用海量数据,精准提升法律监督质效,更好服务社会治理,是新时代检察工作的必答题。检察机关案件管理部门作为检察业务数据的管理部门,更应该深入思考,寻求业务数据资源化、应用管理现代化的最佳路径。

实现业务数据资源化的前提是数据治理。经过全国检察业务应用系统的长期应用,系统内数据体量呈几何倍增,规模级数据资源形成,数据资源化具备了基本条件,但距离资源化应用还存在一定差距。一是部分数据不够准确。由于数据填报质量参差不齐,填报标准把握不一等原因,一些业务数据存在不准确、不真实的问题。二是数据关联不够紧密。目前案件的统计方式是以锁定周期内办案发生量为统计标准,这会导致不同办理阶段的同一案件数据分散统计,较难实现案件关联分析,无法全景式展现案件办理全过程。此外,检察统计数据来源都是结构化数据,有时需要关联法律文书等非结构化数据进行分析,批量检索查询不便利。三是数据标准不够统一。检察业务应用系统、12309中

国检察网等系统产生的内部数据,需要汇聚、分类治理,编制信息资源目录,建立不同主题、专题数据库,比如按照“四大检察”建库,这些都需要通过数据清洗、数据融合等技术手段来实现,数据治理是业务数据资源化的基础。

实现业务数据资源化的关键是数据关联。最大限度发挥检察业务数据资源价值,应主要做好两方面的工作。一是内部数据重在建立关联关系。当前统计报表数据主要是线性、单一案件类别的数据,且因案件诉讼环节不同,一些案件数据分散在不同层级的检察机关,较难直接形成交叉数据、关联数据。在流程监控方面,需要通过获取拟制的法律文书内的信息数据,按照预设的监控规则,自动关联发现办案程序不规范问题;在发现监督办案线索方面,需要通过构建各类型案件之间的关联规则,监测并共享案件线索信息,促进“四大检察”监督融合发展;在发现案件质量问题方面,需要对“一案多件”的案件作批量关联处理,方便启动内部审查,促进司法规范化。二是外部数据重在比对碰撞。检察机关应立足法律监督职责,通过外部数据、内外数据比对碰撞发现监督线索,推动数字检察工作深入开展。比如内蒙古自治区某检察院创建的民事支持起诉监督模型,通过获取全国根治欠薪线索反映平台、全国网上信访信息系统、全区法律援助管理系统的数据并主要进行比

撞,按照欠薪人数、金额筛选出申请法律援助、提出申诉的案件,再与人社局门数据比对,核查劳动者是否属于“农民工”等弱势群体,在征求劳动者起诉意愿基础上,对符合支持起诉条件的依法支持起诉。再比如,某检察院获取当地公安办案平台近两年相关数据,与检察业务数据比对碰撞,发现司法救助线索4条,向金融机构制发社会治理类检察建议3份。此外,通过发现法院三家办案数据的定期比对,还能发现关键案卡错漏填、办“凑数”案等反管理问题,有效提升案件监管水平。

实现业务数据资源化的重点是服务办案和管理。业务数据资源化的目标是提升检察办案质效,强化检察业务管理,让人民群众更好更快感受到公平正义。首先,获取内外部数据,要以法律监督为导向。数据汇聚、获取不能“大而全”,应在盘活内部数据基础上,有针对性拓展外部数据资源。重点获取法律监督必需、亟需的数据。省级检察院还要强化顶层设计,通过建立长效机制来获取数据。其次,数据治理及分析,要以服务司法办案为导向。数据治理是为了方便开展数据检索、比对、碰撞和分析,结果运用是为了高质效司法办案和科学管理检察业务,因此,重点要围绕反映办案力度、质量、效率、效果的数据开展,确保让数据“活”起来,真正释放数据价值。最后,数据共享及管理,要统筹兼顾司法办案与数据安全。数据资源化

是为了更好实现数据共享共用,而数据作为特殊的生产要素,具有可复制、可流动、可共享等特性,应当建立数据资源共享平台,实现数据分级分类管理应用,依申请审批式地有序开放数据资源,确保数据安全。

实现业务数据资源化的保障是技术实现能力。业务数据资源化离不开人工智能、大数据等技术支撑。更为关键的是要统一数据标准,构建数据自动关联、智能分析比对的规则体系,设计可视化数据分析结果输出模式,推送监督办案线索至办案部门。这对技术与业务融合、技术实现和保障能力提出了新的更高要求。一要找准需求。业务主导是检察业务数据资源化的第一属性。技术部门应在充分征求了解办案部门、案管部门业务需求的基础上,探索研究业务数据资源化的技术实现路径。二要集中建设。各省级院要加强顶层设计,统一建设数据获取、数据存储、数据治理、数据分析、数据共享、数据管理、数据安全等多功能于一体的数据资源化管理平台,使数据资源管理、使用更加集约高效。三要不断完善。业务数据资源化的技术实现是一项系统工程,不可能一蹴而就,应在科学设计的基础上,丰富实践运用并不断优化完善。

(作者分别为内蒙古自治区人民检察院案件管理办公室主任、内蒙古自治区人民检察院案件管理办公室检察官助理)

## 检察长眼中的数字检察

□湖南省长沙市人民检察院 检察长 余湘文

数字检察是检察工作提质增效、转型升级的重要抓手,是实现以检察工作现代化服务中国式现代化的强劲引擎。作为湖南省长沙市检察院的派出院,长沙市星城地区检察院坚持检察大数据战略,能动履职,主动探索,将大数据赋能社区矫正巡回检察作为“一把手工程”,以“高质效办好每一个案件”为价值追求,积极加强网络、平台、数据、应用、模型五大融合建设,全力推动新时代法律监督模式变革。

立足社区矫正检察职能和目标,明确数据需求。数字化背景下,数据是与

劳动、资本、土地、知识、技术、管理并列的第七大生产要素,找准数据是做好社区矫正巡回检察工作的重要前提。以社区矫正对象为中心,其在日常工作、生活中将产生大量数据,如通讯、出行、住宿等记录,但并不是所有数据都是有用的数据。什么样的数据才是我们迫切需要的?显然,要立足法律监督职能,立足社区矫正巡回检察工作目标,以问题为导向,从具体场景应用中充分考虑、逐一分析,不断提出需求,描述需求。社区矫正对象单纯的出行、通讯等数据,司法价值不大,但如果是重新犯罪、违反监督管理规定而产生的数据,与执法司法

发生了密不可分的交集,具备明显的司法属性,就成为社区矫正巡回检察中最有价值的数据库。实践中,出入院、社区矫正、交通违法、治安管理等,看守所羁押等数据是我们最迫切的“需求”。如检察官从交通违法数据中发现社区矫正对象存在酒驾记录,这便是有用数据,应充分使用,将其与社区矫正数据中的考核奖惩记录进行比对,可能会发现符合监禁执行的线索。

紧密结合社区矫正检察实际,理顺司法逻辑。司法逻辑是“数字+检察”深度融合的纽带。数据有了司法属性后,就会与其他数据产生或强或弱、或远或近的

司法逻辑关系,法律监督才有着力点与切入点。因此,社区矫正巡回检察中,调取所需的各类数据后,要紧密结合社区矫正检察实际,深入、全面、细致分析数据背后的司法逻辑关系,为找准算法、搭建平台、构造模型、设置场景奠定基础,也有助于提升算力、提高质效。一方面,要考不同类别数据之间的逻辑关系。以星城地区检察院收监执行数据为例,该院检察官在核查后发现社区矫正对象彭某于2022年10月被收监执行,但比对发现,看守所羁押数据、监狱收押数据均无收押此人的记录。经核实,问题在于公安机关未执行收监决定。另一方面,要考

虑同类别数据之间的逻辑关系。如治安管理处罚数据中,检察官通过比对办案单位与社区矫正对象居住地的有关数据,可以得到脱管线索;又如看守所羁押数据中,检察官发现被收监的社区矫正对象在羁押期间仍有签到打卡记录,则可能存在深层次问题。

切实用好人力要素,提升数据效能。人力要素是“数字+检察”的“桥”与“船”,没有人力介入,无论数据是否具备司法属性和司法逻辑,数据仅是数据,没有生命力,无法转化为监督线索。从数据溯源来看,由于人工填报的粗放性、随意性,数据将不可避免地存在瑕疵,如重复、不完整、不真实等。因此,需要检察官运用数字思维,数字方法、数字技术,结合检察业务知识,进行全面核查、细致甄别,将分析对比出来的数据转化为有效的监督线索,为监督办案不断供给“案源”。如根据社区矫正监督模型的提示,检察官发现5名已解矫的社区矫正对象在社区矫正期间醉酒驾驶,但截至目前,无立案、起诉、审判记录。检察官介入后发现,公安机关均已立案,但未录入办案系统,也未移送审查起诉。同时,对于已查清核实的监督线索,也需要检察官做好“后半篇文章”,综合运用到正违法通知书、检察建议书等监督手段,对个案、类案情况予以监督,并不断跟进监督,最终达到社会治理效果。今年以来,该院在开展大数据赋能社区矫正巡回检察中,通过检察官的有力监督,有效挖掘数据效能,监督收监执行21人,移送抗诉线索4条,监督处理司法工作人员5人,监督对社区矫正对象刑事实立案1人。

# 社区矫正巡回检察的“数之思”