

最大限度让数据“活起来、动起来、用起来”

业务数据资源化 大家谈

□徐俊驰 李龙

数据是新的生产要素,是加强检察机关法律监督和改进自身管理的重要资源。最高人民检察院检察长应勇指出,要坚持“眼睛向内”,用好用足内部数据,深度挖掘检察数据价值,让更多“沉睡数据”被充分激活和利用。从工作职责看,案件管理部门负责检察机关内部数据的集中管理,同时与公安和法院系统紧密相连,通过案卷、法律文书等载体汇聚全面的司法办案数据,更应发挥优势,最大限度让数据“活起来、动起来、用起来”,切实做好检察机关数据管理,促进检察机关高效履职。

养成数据思维

数据思维,从根本上来讲,是检察人员基于具体应用场景,建立数据分析框架,用数据提出问题、分析问题、解决问题的方式方法。数字检察战略提出后,检察人员对提高数据敏感度、构建算法和模型、收集数据、研判数据等基本任务形成了普遍共识,积极落实“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。但是,检察人员对检察机关内部数据的整理和应用在某些方面还存在跟不上、不适应的情况,

思想上还需突出强调解决问题的目标导向。因此,检察机关既要加强与其它单位的数据交换、互联互通,也要加强对自身掌握数据的深度挖掘;既要重视以案件办理部门为主,深化法律监督,也要重视以案件管理部门为枢纽,强化日常监管,尤其是对业务流程关键环节、重点领域、突出方面的监控;既要持续强化卷宗完整、案卡准确,也要加快完善各地各条线子系统的贯通,如此才能更好地落实“眼睛向内”,以数据驱动检察业务高质量发展。

强化数据整合

数据资源化的首要前提是大量甚至是全量数据的汇聚。全国检察业务应用系统自上线以来,四级检察院在“一张网”上办案,基本实现了业务数据化。目前,该系统已经集办案、管理、统计、智能辅助、知识服务、数据共享、大数据应用、政法互联等功能于一体,为案件管理部门获取办案信息,实现对业务流程监控、质量评查和数据核查等工作提供了重要前提。但当前,全国检察业务应用系统所汇聚的司法办案数据还未达到全量,较为现实的是网上信访信息系统与全国检察业务应用系统相对独立运行,全国检察业务应用系统无法自动抓取网上信访信息系统办理、回复的信访案件数据,人工二次录入又会造成数据不准确。对此,应加强顶层统筹,尽快打破子系统间的屏障,对全国检察业务应用系统以外的各地各条线

数据端口、应用系统、功能模块等进行整合,化繁为简、合多为一,统一在一个入口下办理业务、在一个系统中汇聚数据,真正将全量司法办案数据汇聚至全国检察业务应用系统中,这是数据资源化需要解决的关键任务。

提升数据质量

数据准确是实现数据资源化的坚实基础。司法办案数据是检察业务工作发展的“晴雨表”,其真实准确是开展数据分析研判、把握办案趋势、服务领导决策的基本前提,更是高质量办好每一个案件的具体体现。实践中,数据质量在一定程度上还存在一些问题,如数据造假、办“凑数案”“注水案”,选择性司法和技术性规避等“反管理”问题。数据质量出了问题,数据资源化就犹如“空中楼阁”。因此,各级检察机关必须坚决纠正“反管理”问题,决不能为了数据好看而采取“反管理”形式,甚至弄虚作假。案件管理部门应该真实、准确、及时地填录案卡,减少缺失数据、错误数据和不可用数据,从源头上提升数据质量。案件管理部门借力相关智能化辅助工具,加大数据真实性、准确性的日常核查,严防数据质量不高,尤其是反复出错、故意瞒报虚报、弄虚作假的情况出现。同时,可加强与业务部门的沟通协作,通过共同指导数据填录、联合开展专项数据检查等形式,及时纠正瑕疵数据,确保数据填录的准确性,有效提升数据质量。

改进数据模型

随着检察业务的持续深入开展,产生、积累的数据越来越多,但这些海量的原始数据往往是存在于各个数据端口、应用系统、功能模块的“数据孤岛”中。数据资源化的目的就是要充分地利用数据,最大限度地发挥其作用价值。构建数据模型是实现数据资源化的重要突破口,通过数据模型,整合全国检察业务应用系统中分散、孤立的数据,进行分析计算、碰撞比对,自动生成新的数据,并以可视化方式予以呈现,使用户能够快速有效地获取数据中有价值的信息,从而作出准确有效的决策。案件管理部门应立足业务监管职责,借助算法和模型的研用来服务监管职能的履行,并间接推动业务部门更好履行法律监督职责,从整体上落实数字检察战略的目标要求。以涉财判项错误监督模型为例,案件管理部门以拒不执行法院判决个案监管为切入点,及时总结涉财案件类案特征,梳理其中反映的问题领域、点位,可能涉及的数据来源、内容等,将其转化为监管规则,嵌入案件管理相关平台,实现相关内容的自动校验、推送、提醒,为业务部门提供相应的监督线索。通过构建监管数据模型,有针对性地调动、利用原始数据,让数据真正“活起来、动起来、用起来”,为推进国家治理体系和治理能力现代化贡献检察业务数据的力量和智慧。

(作者分别为四川省人民检察院案件管理办公室主任、四川省人民检察院案件管理办公室检察官助理)

数字建模的实践起步和未来愿景

□陈岑

数字检察通过大数据、人工智能等技术,推动检察工作高质量发展。模型作为数字检察的重要工具和抓手,从长远发展看,应从统筹管理、技术加持、集约发展等方面推动相关模型的选育构建、应用推广,不断拓展监督的广度、深度和精准度。

一是数字建模规范化。据统计,目前全国检察机关研发运用的模型达6000余个。由于模型建设标准不统一,有些模型未能有效找准切口,场景应用较为模糊,逻辑不严谨。数字检察工作必须做到规范化,尤其是在模型建设方面。因此,最高检、省级检察院要正确处理顶层设计与基层创造之间的关系。比如,应加强统筹管理与指导,制定数字检察工作规则,出台数字检察工作机制;积极推广可复制的大数据法律监督模型,定期开展模型竞赛并发布数字检察典型案例,发挥指引作用;出台模型管理办法,明确建模的目标与要求,规范模型的建设、管理、评审、推广;出台数字检察办案规则、线索管理办法,明确数字检察履职边界,规范线索管理;科学合理地将数字检察融入业务指标,健全数字检察配套制度机制。

二是数字建模智能化。当前,各地检察机关都在进行大数据法律监督模型的探索、分享、推广和应用,依托模型进行数据碰撞,发现监督线索并推动解决执法司法、社会治理领域问题。但检察机关还面临更为复杂、更为隐蔽的治理难题,这些问题难以通过简单的模型或数据叠加进行有效处理,必须借助更

加智能化的模型对数据进行深度运用,挖掘数据要素潜能,实现法律监督的智能化。

因此,要学会运用“思维链”构建智慧模型架构。“思维链”是大模型中的术语,应用到数字检察领域中,就是将案件处理过程及社会治理问题分解为多个环节或阶段的“子模型”“中间模型”,运用“子模型”“中间模型”等智慧模型架构来辅助检察机关应对复杂问题治理。同时,运用恰当的分析方法,如态势分析法、外部环境分析法等,更好地提炼要素特征和规律,为数字检察模型的建设提供基础支持和引导,提高建模的精准度和效率。

三是数字建模集约化。数字建模的集约化,重点是要解决模型相对独立、封闭等现象,实现低投入、高产出。检察机关可从三个角度开展数字模型集约化工作。一是从上级部署的专项工作以及党委政府中心工作角度考虑模型。以矿山安全领域治理为例,检察机关针对未依法审批、设备保障不符合规定、未建立健全事故隐患排查制度等情形建立数字模型,通过不同阶段模型应用之间的聚合与连锁反应,实现深层次治理。二是从职能角度考虑。检察机关各业务条线将建立的“小切口”模型组合成更加系统、更加综合的“四大检察”模型群,实现类型化综合治理。以虚假诉讼法律监督模型为例,具体监督点包括虚假劳动报酬、骗取保险理赔款、民间借贷等,将这些涉及不同业务条线的监督点放在一块,聚合成模型群,将监督问题真正解决到位。三是从工作领域角度考虑。当前,检察机关在环境保护、卫

生健康、国有土地、国有财产保护等领域分别建立了模型,通过把不同领域的模型聚合在一起,一些数据壁垒、隐性问题往往可以被攻破。以广东省清远市检察院研发的涉矿类案监督模型为例,该模型将非法占用农用地刑事立案监督、遗漏处罚事项行政检察、偷税漏税公益诉讼检察等模型聚合在涉矿领域,监督公安刑事立案、督促税务部门追缴税款1.2亿余元,将排查出的公职人员涉嫌犯罪线索移送纪检监察机关,推动了矿产行业深层次治理。

四是数字建模泛化。大数据法律监督模型应用的广泛性是指模型所具有的复制性、可推广性。具体体现在两个方面:一是知识的生产,指通过数字检察生产出新的知识元素;二是知识的形态,指通过数字检察能够形成知识的排列组合。

检察机关在挖掘监督点及归纳要素特征的过程中,不能着眼于某一地域和已有数据,要从全国、全省范围内的数据量中进行发掘。只有获取的数据体量充足,在知识生产与知识形态构建过程中,数据之间才能不断组合并产生新的监督点,进而对数据进行预警式的动态分析和监测,发现潜在的隐患和风险,逐渐形成对数字检察知识的更新迭代,实现社会的有效治理,完成从“管人”向“管数据”的转变。

五是数字建模业务化。数据在模型建设中起着至关重要的作用,具体包括结构化数据和非结构化数据。结构化数据虽然有不同的格式,但处理起来相对容易,格式也便于统一;非结构化数据如文本、图像、音频和视频等,这些数据通常难以用固定的结构化方法来

查询,往往需要使用特定工具和技术进行处理和分析。在数字检察实践中,检察人员需更加注重非结构化数据的使用。但有的检察人员将非结构化数据的使用交由信息技术部门处理,实践中,检察技术人员往往会被漏案件中的关键要素。检察人员作为案件的直接承办人,较为明晰案件办理的一般性规律,了解如何集纳数据,如何选取要素搭建模型等,因此,相较于检察技术人员而言,由业务部门检察官开展要素提炼会更加全面有效。

六是数字建模多元化。检察技术人员运用现代科技手段进行数据挖掘、碰撞、画像及穿透,如利用机器学习、神经网络、自然语言处理技术,对相关文本、相关数据进行语义分析和关键词提取,以此自动识别案件中的关键信息,辅助检察官进行案件分析和定性评估。同时,利用图像识别技术对涉案证据进行自动识别和分析,提高证据的准确性和核查效率,并引入解释性强的机器学习算法,提高模型的可解释性和可信度。数字检察的核心是业务主导下的技术应用,关键在于实现检察业务和技术创新的深度融合,要充分发挥技术的支撑作用,解决业务与技术脱节的问题。实践中,可采用业务人员与技术的培养路径,真正实现以数字赋能检察业务提升。数字检察的发展需要也将涌现一大批“数字检察官”,需注重“数字检察官”的培养,使其不仅具备较强的业务规则梳理能力和监督模型构建能力,还具备运用大数据发现监督线索及案件审查、办理、指导能力。

(作者为广东省清远市人民检察院检察长)

检察长眼中的数字检察

□湖北省黄石市人民检察院检察长 项金桥

迈入数字时代,湖北省黄石市检察机关坚持以数字战略助推检察工作现代化,持续更新数字理念、夯实数字基础、探索机制融合,强化数字保障,逐步激活数字办案的新动能。

数字检察需加强“四大检察”的要素融合

检察机关依托大数据赋能法律监督,需围绕“四大检察”做实业务要素解构和规则提炼,将业务需求细分至计算机可识别、可运算的语言。其中,刑事检察的数字化解构应当以刑法罪名为类型,结合地域数据分析提炼。在关注类罪、涉案主体、行业领域和行为模式等维度的基础上,按照主体、客体、时空等要素分类拆解,对应至结构化数据表库中。业务规则的提炼重点要落在诉讼监督上,既抓住司法办案关键环节的共性问题,又注重本地常见多发犯罪领域的个性问题,通过大数据筛查、比对、碰撞,发现类案监督线索。针对本地常见多发案件,运用大数据精准分析检察办案数据、发案特点及趋势,及时发现多发、高发犯罪背后隐匿的深层次

社会治理问题,精准发现类案背后的系统性、链条性治理漏洞,通过检察监督督促相关部门健全机制、堵塞漏洞,促进源头治理、前端治理。民事检察,行政检察的数字模型构建应在充分考虑经济、文化、历史和地域等差异性因素的基础上做好深层次监督与溯源治理。如可根据不同地方、不同行业领域,将虚假诉讼法律监督模型细分为民间借贷虚假诉讼监督模型、离婚纠纷虚假诉讼监督模型、建设工程领域虚假诉讼监督模型等。同时,随着经济社会发展,以物抵债、托盘交易、非典型担保、新就业形态劳动争议等新型案件层出不穷,这就要求对民事行政检察领域的数据实现快速清洗和可视化管理。公益诉讼检察要把握“公益侵害性”和“行政违法性”两个关键要素,对行政机关的履职做到督促不替代,紧盯行政机关违法行为行使职权或者不作为的行为违法性,从跨部门融合履职、不同执法主体履职权限等角度找准切入点,提炼出需要碰撞的数据及监督规则。对于其他检察业务,要提升数据资源化运用。如控告申诉检察要常态化对控中窗口的涉访涉诉、涉民生保障、涉重点群体领域、涉矛盾风险化解

等数据开展趋势分析、概率统计、科学研判,主动融入法治数字建设。

数据发挥实效的关键在于整合

数字检察应用本身就是“数据+规则”的内涵集合,是数据客体与应用规则相结合的整体概念,因此,数据整合是关键。检察人员只有来自不同系统或数据源的数据进行合并、清洗和归类,使其形成一致有序的数据结构,才能发挥数据价值。以黄石市检察机关数据资源管理平台为例,该平台汇聚了全市核心办案结构化数据及相关法律文书,具体包括检察业务数据、检察管理数据、政务数据、政法数据、互联网数据等各类基础专题类数据,以及电子化档案、文书卷宗及互联网公开的裁判文书、行政法律文书等非结构化数据。通过协商、前置机制对接、资源门户申请等方式,检察机关实现与政法委、公安、法院、人力资源和社会保障、民政等多部门的数据共享交换,基本实现了数据的物理集中,达到本地区检察业务数据共享的整体目标。同时,将来自不同系统或数据源的数据进行合并、清洗和归类,使其形成一致有序的数据结构,实

□吴军 王飞

当前,数字检察工作已全面推开,全国各级检察机关开展数字检察工作的发力点,较多聚焦于研发大数据法律监督模型,运用大数据、区块链、卫星遥感等技术手段,发现数据异常,筛查类案线索,移送相关业务部门落地成案,推动法律监督提质增效。但笔者认为,数字检察要实现高质量发展,还需从多维度进行思考和谋划。

第一维度:实现与党和国家发展大局的深度融合。面对数字时代和社会治理现代化的新要求,传统“等米下锅”式的被动式监督,会让检察机关在服务经济发展大局、优化营商环境等方面产生不足。数字检察是检察机关主动融入党和国家发展大局,顺应数字时代发展形势,提升法律监督能力,推进检察工作现代化的必然要求,更是检察机关贯彻落实全面依法治国战略的实践创新,是检察机关服务大局、为民司法的新引擎、新驱动。检察机关要以更好服务保障经济高质量发展为目标,打造“优化营商环境·护航民营企业”大数据服务中心,建立企业合规数据库,打通政务数据,扩展企业经营异常、不良贷款等外部数据,对内辅助司法办案,对外服务企业发展,通过数字化手段提高法律监督质效,为市场经济提供更加稳定可靠的法治保障。

第二维度:实现与检察机关法律监督属性的深度融合。数字检察应紧扣检察机关法律监督属性,立足法律监督主责主业,引领办案人员更新办案理念、转变传统的办案思维,从个案“小切口”发现类案的逻辑规则和特征要素。数字检察部门应发挥“产品经理”作用,将业务部门提供的类案监督点和监督规则梳理形成法律语言,协调技术部门将其转化为机器语言,通过算法统筹数据和算力,从海量数据中发现类案线索,最终落地成案。比如,山东省日照市检察院研发的通过在押人员生活缴费发现监管民警违法违纪类案监督模型,基于对相关数据的对比分析,发现看守所监管民警违反规定收取索要在押人员亲友微信红包或转账,私自为在押人员传递信件、捎带违禁品等违法违规问题线索,进而督促相关部门完善机制堵塞漏洞,预防此类案件发生。大数据法律监督工作不能脱离法律监督职能,不能超越代庖替代行政工作、主管部门常态化监管执法,这一点需格外关注。

第三维度:实现由“案”到“治”的深度融合。数字时代,违法犯罪形态悄然变化,犯罪逐渐呈现网络化、隐蔽化、智能化、国际化趋势,这就要求检察官应围绕服务大局的热点焦点、执法司法的突出问题、社会治理薄弱地带和公共利益弱项短板等内容展开监督,消除监管的“空白点”,实现由被动向主动、由个案向类案、由办理向治理的转变,以“我管”促“都管”,通过督促相关部门履职,协力破解社会治理深层次问题,推动构建数字检察监督的全新模式。比如,山东省单县检察院针对司法救助案件线索来源少、救助不及时问题,构建国家司法救助线索筛查类案监督模型,并与残联、退役军人事务局、乡村振兴局、妇联分别会签文件,定期召开座谈会,深化落实信息共享、共同帮扶、定期联合回访等机制,凝聚司法救助合力。

第四维度:实现与“四大检察”融合履职的深度融合。大数据赋能法律监督是对传统检察机关法律监督方式的重塑性变革,需要内部统筹协调,各部门积极配合,只有融合履职,形成办案合力,才能适应新时代法律监督工作的新要求。基层检察院应当改变传统办案手段和办案模式,重视向数字检察要战斗力、向融合履职要高质量。通过数字思维、大数据技术,总结归纳出异常案件的特征,增强各类案件线索摸排、调查核实、收集固定证据的能力水平,提升法律监督的精准度。同时,还应建立融合办案团队,实现跨部门一体化、集成化作战,最大限度发挥大数据赋能法律监督的实战实效。

第五维度:实现与信息技术的深度融合。只有实现信息技术与检察业务的深度融合,数字检察工作才有生命力和可持续性。数字检察发展的基础离不开检察信息化手段,根本在于坚持一网运行、一网通办、一网赋能、一网运维,夯实数字检察建设的底座。对此,检察机关必须加快推进法治信息化工程建设,以此打通数字检察的基本路径,加大技术支持,为数字检察的应用构建基础、创造条件。比如,以法治信息化工程为依托的全国检察业务应用系统,汇聚了全国检察机关的办案数据,通过激活“沉睡”的数据,让所有案件信息可以碰撞、提炼、挖掘,从而发现新的办案线索或潜在问题;通过进一步打破数据壁垒,与政法平台、行政执法平台进行合理有序对接,实现业务网上协同办理;通过与大数据法律监督平台有效融合,在更大范围内汇聚其他执法司法数据,提升法律监督能力,促进社会共治。

(作者分别为山东省人民检察院数字检察办公室副主任、山东省日照市人民检察院数字检察办公室副主任)

在要素融合中深挖数据应用实效

帮助处理结构化数据,又能辅助检察官开展案件审查、文书校对和知识记忆;通过区块链技术,检察机关可以将涉案数据经过算法加密后存储在区块链上,实现对案件信息的可追溯、不可篡改和去中心化,确保信息的真实、完整。三是兼顾应用与管理,积极构建数字检察安全防护体系和运行维护体系。采取高效便捷的内部数据管理方式,对全国检察业务应用系统中的内部业务数据,应采取市级检察院审批、省级检察院备案方式,做好对数据采集、存储、传输、交换、应用、销毁等全生命周期的技术维护和安全管理,完善加密级数据存储介质和规范管理权限。

二是兼顾当前与长远,突出实用。在推进数字检察工作中,基层检察院要立足需求,搭建集数据中台、模型引入、低代码模型构建、辅助办案及可视化展示等工具平台。同时,从长远发展考虑,探索应用数字新技术,推进数据深度应用并保障数据安全。如引入GPT等较为成熟的AI平台工具,开展OCR识别、语义分析和要素提取,既能

现对涉案人员及羁押人员、社籍人员、机动车信息、行政处罚记录、未成年人教育等数据信息的自由组合、高效检索和可视化展示。

技术支撑要在科学前提下讲实用

并非高天上的新技术或者功能完备、数据齐备的综合应用平台最具实用性,技术的实用性主要体现在技术应当围绕业务规则和场景需要展开。如检察人员在交叉巡回检察工作中,对于监管人员服刑、减刑、考核和奖惩数据表,可通过表格筛选、对比和碰撞功能实现异常线索推送,方便快速核实和机动调整,在快速建模过程中,检察人员数字素养的积累养成是首要的。

技术的支撑要兼顾以下三方面内容。一是兼顾整体与局部。当前数字检察的探索实践已呈现分层体系,最高检、省级检察院更关注模型共性应用,侧重于模型规则的普适调整和上架推广;基层检察院更关注模型的生产应用和落地效果,从本质上看,外地模型引入本地应用,也存在规则调适、数据对接甚至升级优化的过程,这些均属于模



黄石市检察院检察官开展数据建模研讨。

从五个维度探寻数字检察发展动力