

借力新兴技术解决数据共享安全难题

张洪伟 庄启斌

数字化时代,数据驱动、开放共生的特点正在深刻影响着社会不断进行改革创新,随之而来,检察工作也迎来了数字检察时代。推进数字检察工作稳步开展离不开数据的使用,检察监督范围的延伸,所需数据来源广泛,要获取全部的数据资源显然是无法实现的。数据蕴含着丰富的信息资源,数据的不当使用可能危及商业秘密、个人隐私,带来数据安全隐忧。如何平衡数据资源的使用与安全,是数字检察工作需要解决的重要问题。为此,笔者探索隐私计算在数字检察工作中的应用,实现数据的“只用不见”。

一、研究背景

习近平总书记强调“大数据是信息化发展的新阶段”,随着信息技术飞速发展和国家大数据战略深入实施,大数据在经济社会高质量发展中的作用愈加重要、日益凸显。《中共中央关于加强新时代检察机关法律监督工作的意见》要求,加强检察机关信息化、智能化建设,运用大数据、区块链等技术推进公安机关、检察机关、审判机关、司法行政机关等跨部门大数据协同办案。检察机关数字建设工作已经实现从智慧检务向大数据、区块链、人工智能等多种技术融合发展转变,是推动经济社会发展、完善社会治理体系的大趋势。检察机关数字建设聚焦司法体制改革新形势、新要求,为检察机关开展法律监督、参与社会治理、服务中国式现代化开创了全新的“赛道”。

检察大数据思维的核心,是由传统的“数量驱动、个案为主、案卷审查”的个案办理式监督,向“质效导向、类案为主、数据赋能”的类案治理式监督转变,可概括为“个案发现—类案监督—系统治理”。检察机关对办理的个案进行总结,归纳关键要素,之后,通过与政法、政务、社会等所涉数据进行综合分析,得到具有高度相似性的案件线索,从中归纳提炼数据模型进行类案监督,这一系列过程都离不开数据支撑。全国检察业务应用系统虽然拥有较为完备的检察办案数据资源,但检察机关进行大数据法律监督仍需更多的外部数据予以支撑,数据安全直接影响了数字检察工作的纵深发展。如何突破获取数据的瓶颈,唤醒沉睡数据,在数据安全保护与检察监

督需求之间寻求平衡点,是值得高度关注的问题。

二、数字检察工作的数据现状

一是数字检察工作数据来源分类。数据来源可以分为内部数据和外部数据两种类型。其中,内部数据是检察机关自身收集和记录的数据,主要包括案件数据(案件侦查、审判等阶段的各类数据,如犯罪嫌疑人身份信息、案发地点、证据材料、调查过程记录等)、人员数据(与案件相关的人员信息,如证人、犯罪嫌疑人等个人基本信息、犯罪记录、裁判文书、调查报告、决策理由等)等,这部分数据属于检察系统自有数据,大部分是通过办案过程积累而来,存储于全国检察业务应用系统中,并且属于结构化数据,易于统计分析,使用方便。

外部数据是指除内部数据以外,检察机关行使法律监督职能的过程中,用以丰富案件分析和决策支持的、存在于检察业务应用系统之外的数据。这些数据又可以分为公开数据和非公开数据。公开数据包括政府信息公开的数据等,这部分数据公开使用不存在安全隐患。非公开数据包括各单位行使职能过程中收集到的数据,具有一定的敏感性,一旦泄露会对个人、社会、国家造成一定的损害。使用这部分数据需要注重数据真实性、合法性和隐私保护等问题,在合规合法的基础上进行使用,并注重加强数据保护。

二是数据风险分析。检察机关监督范围的拓展,迫切需要大量数据资源作支撑。而数据共享的同时,也增加了数据安全隐患,数据泄露造成的危害也是数据所有者不愿共享数据的重要原因。数字检察工作中所涉及的信息量庞大,所涉数据多为敏感信息,这些信息一旦泄露,将对个人隐私权和社会公共利益产生严重的影响。因此,在数字化检察实践中,保障数据安全显得尤为重要。

保障数据安全需要从多个方面入手,包括制度、人员、技术等,通过寻找数据共享与数据安全的平衡点来实现大数据共享利用的目标。其中,需对数据的存储引起重视,因为其对数据安全的影

响较大。因此,必须规范数据存储的流程和措施,确保数据得到适当的加密,做到安全性可控。在数据的传输过程中,黑客、内部人员和其他恶意行为可能攻破网络进行监

听,从而造成数据信息的盗取或泄露,甚至会对网络中的数据进行截取修改,导致信息扭曲或失真。加强与数据共享相关组织的管理监管力度,以及完善合理的审计制度等措施有助于防止此类事件的发生。共享数据需要做好数据管理,确保数据具有正确的安全标签、权限和访问控制,保证数据的保密性、完整性和可靠性。

三、隐私计算技术在数字检察中的应用

隐私计算(Privacy compute)是指在保护数据本身不对外泄露的前提下实现数据分析计算的一类信息技术集合。在隐私计算的框架下,数据能够实现“只用不见”;在充分保护数据和隐私安全的前提下,实现数据价值的转化和释放。目前主流的隐私计算技术主要分为以多方安全计算为代表的基于密码学的隐私计算技术、人工智能与隐私保护技术融合衍生的技术、以可信执行环境为代表的基于可信硬件的隐私计算技术三大类。

检察机关作为法律监督机关,无论是立案前的线索发现,还是立案后的证据审查都需要大量的数据予以支持。例如,在进行虚假诉讼法律监督过程中,单独一个案件很难发现监督点,如果对大量数据进行分析研判,就会发现案件的共性,进而挖掘案件线索。然而,数据拥有方出于对数据安全考虑,往往不愿意提供数据,为此,隐私计算兴起为我们提供了盘活数据价值的可行性路径。隐私计算作为一种“可用不可见”技术,可以做到

数据使用的安全可控。该技术将数据的安全性托付于模型算法,防止数据遭受人为破坏。

在多方安全计算过程中,数据的拥有方必须保证数据的绝对安全,这就相当于数字检察工作中数据拥有方出于对数据的安全性考虑一样。在完成一项具体的数据分析碰撞过程中,需要多方共同参与,但彼此互不干扰。整个过程中,双方的数据输入相对于其他方而言是完全隐蔽的,数据存在于第三方模型中,并且彼此都不可通过模型倒推对方的数据输入,唯一获取的数据即为通过模型计算出的结果。在多方安全计算技术的支持下,数据的拥有方不必将数据移交或者共享出去,即可释放数据的价值。

在第二类隐私计算技术中,检察机关与其他数据拥有方作为参与者形成一种联盟,彼此间形成一个全局的初始化数字检察办案模型并分发参与方。通过计算,检察机关与数据拥有方利用可靠网络传播进行模型修改,实现数字检察工作对外部数据的使用,同时能够很好地保护数据安全。比如在检察工作中对民间借贷执行款漏缴个人所得税进行法律监督过程中,检察院需要与法院、税务部门联合搭建漏缴利息税法律监督模型,再通过各自的数据进行模型训练,将获取的利息税被执行人信息、利息税缴纳人员信息上传至模型中进行数据的聚合、分析。

可信执行环境即通过隔离出一个安全的环境来运行程序,保护其中的数据,为数据的分析处理提供安全保障。可信执行环境是基于硬件层级的安全保护,能够有效防止敏感数据泄露,在数字检察工作中,可以运用到对检察业务应用系统产生的数据进行高效安全的分析、研判。

隐私计算在检察工作中的应用可以提升数据安全性和个人隐私保护,同时提升数据共享、案件处理的效率和准确性,为检察工作数字化转型和改进提供支持。

最高人民检察院检察长应勇强调,数字检察、检察技术都要以服务高质效办案为目的,要按照“分头建设、分散储存、关联应用”,在用好自身内部数据的基础上,通过建立机制,有效利用外部数据。由此可见,数据的共享与安全是当下数字检察工作开展的基石,数据的“只用不见”是盘活数据资源、实现数据有效利用的最佳方式。

(作者单位:山东省枣庄市薛城区人民检察院)



薛城区检察院党组书记、检察长张洪伟汇报展示该院数字检察情况。



薛城区检察院干警通过大数据比对分析,查找案件线索。

建立类案监督模型 提升检察监督质效

由于司法实践的复杂性和多样性,“以钱买功”“虚假立功”等异化情形时常出现。2022年以来,山东省枣庄市薛城区检察院通过建立危险驾驶罪立功类案监督模型,精准锁定立功案件监管盲区,并利用数字检察破解虚假立功案件治理难题。

薛城区检察院在办案中发现,近两年来危险驾驶类案件中犯罪嫌疑人举报立功线索明显增多,且部分立功线索来源模糊、随意性强。随后,该院对2020年以来立功案件进行了全面细致审查,发现在李某某危险驾驶案中,存在举报人立功与被举报人自首相矛

盾的情形。

以此案为突破口,薛城区检察院以“危险驾驶罪、立功、被举报人姓名、自首”为关键词建立数据监督模型,依托大数据比对分析,梳理出2020年以来办理的危险驾驶案件中涉及立功线索15件,通过对举报人立功线索与被举报人到案方式进行比对,最终发现有3名被举报人同时被认定了自首。目前,该院已对这3起案件启动监督,并将根据监督结果依法启动审判监督程序或向职务犯罪侦查部门移交线索。

(王辉 李嘉懿)

山东枣庄薛城: 以“数动检行”为指引打造数字检察品牌

一项数字检察成果获全省检察机关创新成果评选“一等奖”,一个数字模型获全省检察机关大数据法律监督优秀模型评选“一等奖”,受邀在全省数字检察培训班上作典型经验介绍,在最高人民法院数字检察专题调研时作汇报演示,数字检察文化获得全省“十佳”……这是山东省枣庄市薛城区检察院在数字检察工作上取得的丰硕成果。近年来,该院新一届党组抢抓机遇,超前作为,培育出以“数动检行”检察文化品牌,全体检察干警在数字检察理念引领下,努力争当新时代“数字工匠”,推动各项检察工作走在



“数动检行”检察文化品牌为数字检察注入活力

“薛城是铁道游击队的故乡和奚仲故里,有着开拓创新、机智灵活的文化底蕴。多年来,我院科技强检的意识和行动不断增强,勇于率先迈出数字检察步伐。”薛城区检察院党组书记、检察长张洪伟在该院“数动检行”检察文化品牌被评为全省“十佳”检察文化品牌的现场汇报中讲到。

2021年,薛城区检察院自主研发的智能检务协助查询系统,实现了各类资格联审环节刑事犯罪信息查询时长从“天数”到“秒数”的量变,带来“只跑一次”到“一次不跑”的重大变化,该做法在全省检察系统得到推广使用。此外,该院在

一起诈骗案件进行审查时,技术人员使用技术手段处理图片,甄别出关键证据上肉眼无法区分的伪造签名痕迹,使案件取得关键性突破,节约了办案时间和资金成本。目前该技术应用已被制作成视频,为全市检察机关提供教学指导。

发挥数字检察新动能让数据“活起来”

在审查一起虚假诉讼案件时,办案人员发现案件当事人通过虚假诉讼手段取得裁判和执行文书,骗取车辆牌照拍卖款。循着该线索,办案人员

从中国裁判文书网调取2016年以来该市民事裁判文书相关数据,以关键词进行检索,排查出同类线索53件,通过开展数据比对分析后,最终确认虚假诉讼案件30件。该院依法监督,法院全部改判,取得了办案的历史性突破。根据该

院移送的职务犯罪线索,全市查办了审判员职务犯罪案件3件3人。

“我们在办案过程中切实尝到了大数据应用的甜头。”承办检察官王冲如是说。该案的成功办理,让检察官真切感受到数字检察的魅力,“数化”万物”

意识深深镌刻在每名干警心中。

2022年6月,全国检察机关数字检察工作会议召开后,薛城区检察院党组决定在全市率先探索数字检察改革,定下“深化学习树牢数字化理念、研讨交流养成数字化思维、逐步应用打造数字化化数字”三步走目标,部署开展数字检察“百日攻坚”,向着“数字赋能监督、监督促进治理”新模式迈出坚实步伐。

在数字化改革先行地浙江学习归来后,该院迅速建设数字检察中心,搭建数字办案平台,构建包括危险驾驶罪立功情形类案监督、司法救助线索筛查、民间借贷利息应纳税所得税公益诉讼法律监督等,涵盖“四大检察”的数字检察模型15个,汇聚各类数据上万条,筛查监督线索300余条,办理各类案件91件。该院还举办全区首届数字检察建模竞赛,真正让数据“活起来”。

扬帆数据蓝海的检察“后浪”

“数字检察应当以业务为主导,实用才是关键。”“模型搭建后,数据的汇集仍然是一个难点。”青年干警在薛城区检察院奚仲法学社讨论得热火朝天。这种数字检察专题研讨会,该院每月至少开展一次。

薛城区检察院致力于发挥青年干警创新优势,将青年干警培养成数字检察改革的主力军。模型搭建,现场推演、数据应用各环节,全部由青年干警作为主力参与其中。同时,该院开设数字检察论坛,创办数字检察特刊,推动理论创新、文化引领,数字思维一体提升。年轻检察官逐渐养成“案前找数据,办案用数据,案后汇数据”的习惯和能力。

点点微光,汇聚成炬。薛城区检察院“数动检行”检察文化品牌,以其独具匠心的创新实践,涵养全院干警敢为人先的创新意识,激发了事不避难、主动出击的“亮剑精神”,磨砺出言必行、行则必成的执行能力,为探索大数据赋能检察监督新路径提供薛检方案,贡献薛检智慧。

(孙璇 孙开磊)

大数据赋能破解社区矫正对象违规外出监管难题

“没想到违反了有关规定偷偷去外地游玩,也能被查到。”当山东省枣庄市薛城区检察院检察官找到违规外出的社区矫正对象褚某某时,他感慨地说。

2022年7月28日,法院以帮助信息网络犯罪活动罪判处褚某某有期徒刑八个月,缓刑一年,缓刑期间依法实行社区矫正。然而,耐不住寂寞的褚某某在接受社区矫正期间,偷偷去外地游玩。褚某某以为只要自己不说,监管部门就无从查起。殊不知,检察官运用大数据社区矫正法律监督模型,对社区矫正人员的出行轨迹和请假等信息进行分析比对后,发现了褚某某违规外出线索。随后,针对褚某某违规行为,检察官及时与其谈话,褚某某表示将改正错误,坚决遵守纪

律,用心接受矫正教育。

据了解,针对社区矫正人员违规外出发现难、处理难,社区矫正工作人员数量少等问题,薛城区检察院以大数据思维设计了社区矫正人员违规外出数字监督模型,研判提取外出轨迹、请假信息、处罚情况三大类监督要素,以时间、空间二维角度碰撞对比,实现批量发现、批量办理,破解以往社区矫正人员违规外出监管难题。该监督模型运行一年来,该院在数据筛查中发现27人次疑似违规外出线索,查实18人次属于违规外出。随后,该院依据法律监督职责,督促社区矫正机构完善监管措施。今年6月,该数字监督模型被山东省检察院在全省推广应用。

(吴勇 郭凯)

漏缴个税款问题 得到了解决

“针对案件当事人利息所得漏缴税款的问题,我们运用大数据建立公益诉讼法律监督模型,通过信息比对碰撞,漏缴税款行为就会无所遁形。”近日,山东省枣庄市薛城区检察院检察官王冲介绍。

2022年9月,薛城区检察院在办理金某诉他人民间借贷纠纷民事监督案件时发现,判决生效后,金某通过向法院申请执行的方式获取借贷利息2万余元。但金某收到执行款项后,从未对利息进行个人所得税申报。该案暴露出利息所得缴纳税款存在监管漏洞,导致国有财产流失。为弥补监管漏洞,推动该领域实现系统治理,该院

决定运用大数据思维搭建公益诉讼法律监督模型,筛查出漏缴税款的线索。

薛城区检察院通过分析案件要素,确定模型所需数据;寻找类案共性,确定关键词;构建监督模型,建立“数据比对分析、清洗案件信息和组织数据碰撞”三步走模型运行逻辑,建立起完整的自然人之间执行所得利息应缴纳个人所得税公益诉讼监督模型。随后,该院运用该监督模型成功筛查出40余起漏缴个人所得税案件线索,并将线索移交当地税务部门,推动追回税款30余万元。

(韩荣新 殷正)