



大数据 侦查

Big Data Criminal Investigation

王燃 著

清华大学出版社



大数据 侦查

Big Data Criminal Investigation

王燃 著

清华大学出版社
北 京

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

大数据侦查/王燃著. —北京:清华大学出版社,2017

ISBN 978-7-302-45551-6

I. ①大… II. ①王… III. ①数据处理—研究 IV. ①TP274

中国版本图书馆 CIP 数据核字(2016)第 270844 号

责任编辑:刘 晶

封面设计:汉风唐韵

责任校对:宋玉莲

责任印制:李红英

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 刷 者:三河市君旺印务有限公司

装 订 者:三河市新茂装订有限公司

经 销:全国新华书店

开 本:160mm×230mm 印 张:14.5 字 数:196 千字

版 次:2017 年 1 月第 1 版

印 次:2017 年 1 月第 1 次印刷

定 价:49.80 元

产品编号:072567-01

本书出版得到天津大学法学院资助

序

爱丽丝：“请你告诉我该往哪个方向走。”

柴郡猫：“这取决于你要到哪里去。”

爱丽丝：“我并不在乎要到哪里去。”

柴郡猫：“那你走哪条路都没关系。”

爱丽丝解释道：“我只想去任何一个地方。”

柴郡猫：“你一定能够实现这个愿望，只要你走的够远就可以了。”

那还是2014年，我开始研读英国学者舍恩伯格的经典之作《大数据时代：生活、工作与思维的大变革》，被其中的恢宏描述震撼到了。之后又看了中国学者涂子沛的《大数据》等相关著作，进一步被大数据战略、数据革命、数据帝国、数据治国、数据开放、大趋势、大挑战以及大变革等词汇所打动。随后，我开始郑重思考自己所在团队——中国人民大学电子证据研究小组（我们自称“人大团队”），在法学研究方面该不该向大数据法律和司法方向转型。这时，我看到了当时网上风靡的帖子，它以《爱丽丝梦游仙境》的场景为例，讲述了大数据技术中数据挖掘的魅力——任何业务问题都可以转换为数据挖掘问题。我理解，这是大数据时代的寓言。

“人大团队”并不是一个严谨的学术机构，它是由人大法学院、信息学院、信息资源管理学院的师生基于共同的兴趣走到一起形成的。它也有研究平台和实务平台，前者包括人大法学院的证据学研究所、网络犯罪与安全研究中心、知识工程与数据工程教育部重点实验室等；后者包括中国人民大学物证技术鉴定中心、人大法学院证据学实验室等。后来，我们又陆续聘请了公检法纪以及公证、鉴定机构、科研院所等部门朋友参加。逐渐壮大了队伍，形成了覆盖电子证据全行业的规模。有了共同的研究旨趣，“人大团队”做了很多针对电子证据的法律与技术交叉、理论与实务跨界的

工作,在电子证据的法治建设、理论创新、实务推动方面做得颇有声色。“人大团队”并没有名义上的负责人,我的恩师何家弘教授算是“精神领袖”。早在2000年9月,他访问日内瓦国际电信联盟等机构归来,敏锐地决策要认真研究电子证据问题。^{〔1〕}这是“人大团队”面向IT时代的一次布局。事实证明,此次布局是非常有远见的,且相当成功的。

那么,“人大团队”在DT时代该做出什么样的贡献呢?变与不变,就是首先面对的问题。一方面,“大数据泛滥”。许多人“言必大数据”,但真真假假、虚虚实实,浮夸的成分不少。大数据能否支撑一个时代,能否改变社会方方面面,当时尚不明确。即便大数据技术就是时代性的,我国是否需要配套的司法治理、法治建设,也令人疑惑。“人大团队”一旦转型,能否一如既往地形成独特优势,也需要琢磨。另一方面,“法律人不能缺席大数据”。大数据是一座巨大的金矿,法治的阳光不能照耀到是不可能的。2013年美国奥巴马总统(“大数据总统”)将之定义为“未来的新石油”,将“大数据战略”上升为国家意志,声称未来对数据的占有和控制甚至将成为国家核心资产。就国内而言,2011年温州动车事故也开始让人们领略到了社交大数据的威力。当年7月23日20时30分05秒,D301次列车与D3115次列车在温州发生动车组列车追尾事故。在专案组成立之前,新浪公司就发布了3286883条关于这起事件的微博;之后,基于700多万条微博制作了视频,从事故现场、寻人、遇难名单、献血现场等多角度展示这次突发事件的真相。至此,我冥冥中受《爱丽丝梦游仙境》柴郡猫说法的启示,决定拓展团队研究范围。

事实证明,这一决策是正确的。大数据发展的潮流是不可抗拒的,大数据法律和司法的改变也是亟需的。中国计算机学会大数据专家委员会在“2013年中国大数据发展白皮书与2014年大数据发展趋势预测”报告中论断,2014年将是大数据从“概念”走向“价值”的元年。2015年后,我国的大数据发展急剧加速:7月,《国务院办公厅关于运用大数据加强对市场主

〔1〕 何家弘主编:《电子证据法研究》,前言部分1~2页,北京,法律出版社,2002。

体服务和监管的若干意见》发布,提出运用大数据加强对市场主体的服务和监管;8月,国务院发布《促进大数据发展行动纲要》,推动各行业大数据发展,强调数据资源共享开放;10月,党的十八届五中全会明确提出“国家大数据战略”;今年3月,“十三五”规划纲要出台,再次强调国家大数据战略。司法系统也开始加入大数据的时代潮流。仅就公开的新闻报道来看,今年下半年中央政法委孟建柱书记、最高人民检察院曹建明检察长、李如林副检察长等纷纷率团调研贵州大数据交易中心,为“大数据+司法”进行布局。尤其是2016年10月21日,上午高检院召开了大数据应用研究会,下午中政委请马云给全国政法队伍讲授大数据等科技创新在社会治理中的运用。大数据法律共同体已经全面行动。

“人大团队”较早地转向大数据法律研究,也就是打开了另外一扇窗。2015年5月13日,我第一次受邀给全国军队保卫部门讲授《大数据侦查与大数据证据》,获得了良好的赞誉,尽管当时的认识还不够深入。此后,我陆续以《大数据时代网络安全问题与挑战》《大数据推动检察办案变革》《大数据在检察办案中的运用》《大数据在纪律审查中的运用》等为题开讲,得到了越来越多的认同,在此过程中也与一线办案同志交流了宝贵经验。实务中,我们积极改造所熟悉的手机取证业务,拓展基于大数据取证的司法鉴定工作,并在个案中获得成功。我们还敏锐地发现,几乎所有的大数据公司都通过手机APP,收集广大用户位置等信息,而这一现象将会极大地改变现有的侦查制度、证据制度和权利保障制度。2015年12月,每年一度的网络犯罪高峰论坛召开,我代表团队就“大数据的证据价值、侦查模式与权利保障”发言,以丰富的实践案例和直观的技术图示吸引了场内外广大专家学者的热议。

今天看来,“万物皆数据”,“数据司法是未来科技司法的主方向”,“司法人员将越来越离不开大数据引领”等言论,越来越成为新的共识。这些规律我们较早地感受到了,也作了一些有益的探索与推动。“人大团队”在许多场合都呼吁,我国的网络安全、反贪侦查、纪律审查、食安执法、网信执法、文化执法等工作都应加强大数据的运用,深化同大数据公司的合作,同

时有效规制大数据的安全与共享问题,并能够给出具体的方案。

大数据法律与司法问题归根到底是下一代年轻人的舞台。在这一进程中,“人大团队”很多年轻人开始持续发力。谢君泽老师在挖掘电子文档痕迹方面有着独到的认识,他基于对批量产生的文档痕迹进行分析,成功地协助查办了国家审计署审计华润公司煤矿并购的泄密案件、天津港8·12爆炸案之安评部门渎职犯罪案件等。我去微软中国公司交流时,方得知这可以叫做“大痕迹数据”。君泽虽不是我名下的学生,但却是辅助我时间最长、最得力的助手和骨干,现已名声在外,前途无量。徐菲、张杨杨、郭树正同学很早就配合我对电子定位技术进行研究,探索收集APP背后的IP地址、MAC地址等大数据的方法,并已日臻成熟;周迪、吕宏庆同学擅长互联网数据挖掘,对网络舆情分析、数据画像及数据碰撞等率先探索;陈泽鸿、张洪铭同学积极试用人大信息学院开发的“时事探针”平台,试探性地绘制了我国的反腐败指数图、网络犯罪指数图;张艺贞、黄蓉同学较早借鉴国外“OPEN DATA”机制,对国内公开数据库如何归整利用进行实验;胡聪同学运筹帷幄,组织团队对BAT公司调研,推动网信部门和检察部门建立大数据公司有效协查调证机制;王耀同学撰写《职务犯罪侦查的大数据模式初探》一文,展示了反贪工作中借助大数据的现实与前景……这样的优秀学生很多,他们以自己的方式在感受大数据时代的脉搏。

此外,“人大团队”特别注意与“外面”的大数据专家合作。“外脑”的指导对于我们开展研究起到了关键性的作用。例如,人大信息学院院长文继荣教授曾经长期任职于微软公司,我们多次登门拜访求教大数据知识,文教授不吝解惑,并无偿向我们提供了“时事探针”应用平台,还根据我们的需要特意对中国裁判文书网的海量裁判文书进行大数据分析;人大公共决策实验室王克平主任多次为我们提供最先进的大数据可视化展示实验室,不厌其烦地展示大数据在公共决策、司法办案中的运用;人大信息资源管理学院的钱毅等教授也伸出援手,协助我们成功申报国家社科基金项目“大数据时代电子文件的证据规则与管理法制建设研究”,促成了一个跨越法学与电子文件管理学的大数据研究机会。中国科学院高能物理研究所

的许榕生教授、香港大学 K. P. Chow 教授也不吝赐教,分享了他们在大数据分析及预测方面的宝贵经验。还要特别感谢来自我挂职的检察系统,以及检察行业的朋友。他们让我们看到了大数据与检察工作、大数据平台建设、大数据预防、大数据初查、大数据侦查、大数据管理、大数据挖掘、大数据碰撞、大数据画像等鲜活例子,也讲授了他们在实务中积累的宝贵经验。大数据转型研究之路上,这样的同道者,我们有很多很多,铭记于心。

王燃博士也是“人大团队”一员,是最值得称赞的大数据法律制度探索者。我依稀记得她初到人大法学院证据学教研室的情景。那一二年级共有五位法学硕士,她看起来话不多,抽签师从我的恩师何家弘教授(跟我同辈呢)。不过,我也给她上课,带着她做项目。硕士两年、博士三年下来,她给我的印象——不是最聪明的学生,但却蛮有智慧,更是执行力超强。马云说过,大数据时代电脑一定比人类聪明,但人类永远比电脑有智慧。王燃是不是“人大团队”中的有智者呢?天知道,反正她选择了大数据法律和司法作为研究方向。

忘了是什么时候,她征询我关于博士研究方向的建议。我可能随口说了大数据法律问题研究很有前景。那时我的内心想法是,团队必须研究大数据法律问题,但这个主导者可能未必是她。结果她认真了,很快拿出了文献综述和写作提纲。而这个题目对于她而言,显然是有相当难度的。她既没有技术背景,也对实务不甚了解,还不了解海外发展情况。谁知道她会怎么切入研究?她会不会做出成果?

她的智慧就是“认定了就做”。她挤出时间到北京市检察院挂职,尽快了解实务;她访学台湾地区,了解境外情况;她更瞅准时机向各位老师求教,博采众长;她还虚心向法律硕士的师弟师妹们学习手机取证、大数据分析等经验,弥补了自己技术盲的短板。我记得博士论文开题时,她拿出了一份“不太好”的写作提纲。导师组建议重新梳理另起炉灶,而我直接提议她集中研究当时已经热兴的大数据侦查,写透大数据侦查的思维、模式、措施、制度等基础问题。没想到,半年后她真的交出了论文稿。当然,她也付出了身心交瘁的代价,她经常跟熟悉的同学开玩笑说最后悔读博士

了,弄得一脸痘。其实,她博士论文答辩通过时满是喜悦,在场的每个人都能够感受到她的心情。几个月后,她便将博士论文修改完善出版,这也是执行力强的明证。

当前我国政法系统正积极向大数据技术靠拢、向大数据战略转型。这时收到她《大数据侦查》专著文稿,我也非常欣慰。“人大团队”终于有成员拿出了大数据法律的第一本专著,这应该也是国内的第一本大数据侦查论著。我想,这就像我2004年出版《中国电子证据立法研究》专著一样,走出第一步就意味着良好的学术开端。我相信,她还会推出诸如《大数据证据》《大数据权利法律保护》之类的“几部曲”。据我了解,她的论文《大数据时代侦查模式的变革及其法律问题研究》荣获了第11届中国法学青年论坛主题征文一等奖,她以“大数据侦查与大数据证据”开启了天津大学的“北洋法学学术沙龙”第一讲,她还受邀给全国检察机关第一次大数据专班主讲“大数据证据”。崭露头角的她,会在这条道路上走好走远!

以我的学术眼光,王燃博士的《大数据侦查》一书具有相当的创新性:一是概念的全面创新。她构建起大数据侦查较为完整的框架,包括概念、思维、模式、方法及相关制度构建等。二是重要观点具有前瞻性。本书很多观点是在博士论文中表达的,当时提出的很多观点现在看来具有相当的前瞻性,并已逐渐被证实。例如,书中揭示了大数据热潮下的一些思维误区;强调大数据相关思维和预测思维在侦查中的广泛运用前景,尤其是预测性,必将推动事后侦查向事前侦查、预测侦查转型;提出大数据侦查的模式,强调从数据空间去寻找突破点;提出大数据搜索、大数据碰撞、大数据画像、大数据挖掘、犯罪热点分析、犯罪网络分析、大数据公司取证这几种大数据侦查的具体方法,已经越来越为实务部门所开发运用;强调大数据侦查在发展技术、应用的同时,要注意其所带来的法律问题以及对传统法律原理、规则的冲击,应当对大数据侦查进行一定的程序规制。诚然,这本书也难免有幼稚和错误之处,这有赖于读者们的慧眼识别。

马云还说过,“整个大数据时代最重要的事情,是要做到‘事前诸葛亮’,就是要有预防机制”。《大数据侦查》一书在某种程度上也是“事前诸葛

亮”。王燃博士是不是在两年前就预测到了“大数据+司法”在今天的热络呢？是不是也昭示着“大数据×司法”在未来的突起呢？

大数据时代是充满无限生机的时代，也是一切都有可能时代。王燃博士出版《大数据侦查》为人们提供了一个“柴郡猫”智慧的小样本。同时，本书的出版也为“人大团队”的大数据之行留下了一个印记。我相信，这本书开卷有益。我期望，“人大团队”在DT时代做出新的华丽转身。

刘品新

2016年10月22日写于拉萨

自序

本书的设想最早形成于2014年11月。尽管当时我国官方尚未提出大数据战略,但大数据技术已经在电子商务、互联网、金融等先驱领域开始运用,国际上也有很多国家相继开启了“大数据革命”。欧盟委员会早在2010年就提出了“欧盟开放数据战略”;联合国推出了“全球脉动”(Global Pulse)计划,建立世界范围内的预警机制。美国、日本、英国、法国、韩国、新加坡、印度等国都将大数据纳入了国家发展计划。^{〔1〕}彼时,笔者开始意识到大数据巨大的发展潜力和前景,并考虑在侦查领域、司法领域推广大数据战略的可能性。结合我国当时的信息化侦查水平、网络侦查制度、电子取证等技术的运用,又了解了其他国家大数据在司法领域的运用情况,如在美国刑事侦查中“大数据预测警务”技术(predictive policing),美国民事诉讼电子证据开示中的“大数据智能检索”技术(predictive coding)等。笔者认为,大数据在我国的侦查领域将有广阔的运用前景。

近几年大数据的热兴也印证了笔者的想法。各侦查部门纷纷搭建大数据应用平台,发展大数据侦查技法,提出“智慧公安”“科技强检”等口号。但目前实践中各侦查部门的大数据运用尚处于摸索阶段,并没有形成统一制度,相关技术方法的运用尚不成熟,相关权利、程序缺乏法律保障。针对侦查实务中大数据运用的蓬勃之景,笔者以前瞻性的视角提出“大数据侦查”这一全新概念,对大数据侦查的内涵、特征、思维方式、技术方法进行了归纳和总结。另外,在发展大数据侦查的过程中,大数据本身的技术、思维特征也会不可避免地对一些传统侦查程序造成影响,对公民的相关权利造

〔1〕 参见工业和信息化部电信研究院:《大数据白皮书》(非出版物),2014年5月,载工业和信息化部网, <http://www.miit.gov.cn/n1146312/n1146909/n1146991/n1648536/c3489505/content.html>, 2016年9月20日访问。

成侵害。基于这些问题,笔者提出大数据侦查的程序规制和权利保障制度,以及数据共享、技术构建、行业规范等相关的配套制度的建设。除第一章导论外,本书共分为五个章节。

首先,关于“大数据”及“大数据侦查”的内涵。大数据包括海量数据集、数据处理技术及数据分析结果这三层含义。大数据不仅仅是海量数据的集合,也是集数据处理、数据分析于一体的技术体系,同时也强调反映事物背后规律的数据分析结果。正确理解大数据的内涵还需要注意,大数据的基础在于数据化;大数据的量多是相对的,对于分析对象来说,达到“样本=总体”的程度即可;大数据的核心价值在于数据背后的规律而非数据本身,而数据规律主要依靠数据挖掘等大数据技术来实现。相比于小数据时代的思维方式,大数据具有全数据、混杂性以及相关性的特征:全数据意指人们完全可以获取某个研究对象的所有数据,不需要再通过抽样调查的方式进行统计;混杂性意指不需要每个数据都精确无误,数据的量多可以抵消部分数据的不准确;相关性则是指大数据颠覆了人类长久以来的因果关系思维,大数据能够快速告诉我们事物之间的相关关系是什么,却无法解释背后的原因。

在此基础上,笔者对大数据侦查的内涵和外延进行界定。从狭义上来说,大数据侦查强调采用大数据技术的侦查行为。大数据侦查是指法定侦查机关针对已发生或尚未发生的犯罪行为,为了查明犯罪事实、抓捕犯罪嫌疑人、预测犯罪等,所采取的一切以大数据技术为核心的相关侦查行为。具体而言,大数据侦查的主体是法定侦查机关,侦查的对象是已经发生或尚未发生的犯罪行为,侦查的目的是查明犯罪事实及预防犯罪活动的发生,侦查的内容是涉及大数据技术的一切侦查行为。从广义上来说,大数据侦查不仅仅指技术层面的侦查措施,而是包括大数据侦查思维、侦查模式、侦查机制等完整体系。相比于传统侦查而言,大数据侦查具有以下特征:侦查空间的数据化,大数据侦查在平行的数据空间中展开,找到与物理空间人、物相对应的数据形式;侦查技术的智能化,大数据本身就集人工智能、计算机等多个学科于一体,数据收集、数据清洗到数据分析的每一个环

节都离不开机器的支持,因此大数据侦查技术必然也具有智能化的色彩;侦查思维的相关性,传统的侦查是一个由果溯因的重构犯罪过程,建立在相关性基础上的大数据侦查改变了这一传统逻辑,直接通过数据运算去发现各要素之间的关系,从而发掘侦查线索。大数据侦查作为一个全新的概念,也需要厘清其与技术侦查、侦查技术、信息化侦查、情报导侦等概念之间的关系。大数据侦查与技术侦查是交叉关系,大数据侦查中对某些数据的收集需要遵守技术侦查的规制;大数据侦查从属于侦查技术的范畴;大数据侦查与传统的信息化侦查、情报导侦之间是传承与发展的关系,大数据侦查建立在信息化侦查、情报导侦的多年发展基础之上,同时又大大推动了二者的发展。在目前的侦查实务中,大数据主要作为线索运用,但不排除大数据在将来会成为一种新的证据形式。总而言之,大数据侦查有利于推动事后侦查向事前侦查转型,被动侦查向主动侦查转型,单线侦查向协作式侦查转型,粗放式侦查向集约式侦查转型,它必将引领未来侦查发展的新方向。

其次,关于大数据侦查的思维特征。笔者结合大数据本身的特征和其在侦查中的实务运用,提出了相关性、整体性和预测性三大特征。相关性思维能够告诉人们事物之间的关联性但不能解释为什么。利用相关性,侦查人员可以找到犯罪现象的关联物,通过关联物来观察犯罪行为本身;还可以通过大数据的相关性分析发现更多隐藏的线索。整体性思维强调大数据时代取证思维的整体性和事实认定的整体性,在整体数据中寻找与案件有关的数据,在整体事实中选取与案件有关的事实。预测性思维则强调对未来时空犯罪活动的预测,包括对人、案及整体犯罪趋势的预测,从而有利于侦查人员合理部署侦查资源,防患于未然。当前,在“大数据热”的氛围中,也容易产生一些思维误区,如认为数据越多越好、数据可以不精确、大数据分析结果一定是正确的、大数据的相关性可以替代因果性、大数据的预测性违背无罪推定原则等。然而,大数据并非是万能的,数据采集中会有偏差,数据结果也会受到人为主观操作影响,大数据还会产生歧视和偏见,数据分析模型也会失灵。另外,大数据侦查的相关性思维特征还会

对传统司法证明原理带来冲击。如何去协调传统侦查思维与大数据侦查思维的碰撞、如何在现有法律框架内去发挥大数据侦查的思维价值,是不得不面对的问题。

在前述基础上,本书归纳了大数据侦查的几种典型模式。在实务中已有的大数据侦查案例基础上,笔者从对象、时间等不同角度将大数据侦查提炼为不同模式。按照侦查对象的不同,大数据侦查可以分为个案分析模式和整体分析模式,前者主要针对具体个案的侦破,后者则面向于整体历史案件的多维度分析。按照时间序列的不同,大数据侦查可以分为回溯型模式和预测型模式,回溯型模式是针对过去已经发生的犯罪行为,而预测型模式则是针对未来未知时空的犯罪,强调对犯罪活动的预测。按照数据形态的不同,大数据侦查可以分为原生数据模式和衍生数据模式,在原生数据模式中,大数据只是作为一种技术、媒介,发挥的是“找数据”功能,不会改变数据的原始状态;而在衍生数据模式中,大数据则对原始数据进行了二次挖掘,发挥的是“分析数据”功能,获取的是新的数据形态。从“数据化”的特征出发,可以将大数据侦查分为“人—数—人”和“案—数—案”模式,前者是指在数据空间找到对应的数据化嫌疑人,后者是指在数据空间找到对应的数据化案件信息,两种模式都遵循着从具体到抽象的过程,大数据在两种模式中都扮演着连接现实空间和数据空间的中介。在传统“由案到人”和“由人到案”的基础上,大数据侦查可以分为“案—数—人”和“人—数—案”两种模式,前者是以案件为中心去找嫌疑人,后者是以嫌疑人为中心去寻找案件事实,它们的共同点就在于通过大数据连接起案件与嫌疑人之间的关系。

再次,本书介绍了实务中常用的几种大数据侦查方法,包括数据搜索、数据碰撞、数据挖掘、数据画像、犯罪网络分析、犯罪热点分析以及大数据公司取证等。数据搜索是较为简单的方法,其原理就是在海量数据库中检索出相关数据,具体包括数据库搜索、互联网搜索和电子数据搜索几种方式。侦查人员要注意发挥大数据智能化检索技术、一键式检索技术。数据

碰撞意指通过多个数据集之间的自动比对来发现相关数据,数据碰撞往往能产生意想不到的效果。常见的数据碰撞类型有话单数据碰撞、银行数据碰撞等。数据挖掘是大数据较核心的技术,包括关联性分析、分类分析、时序分析等多种功能。数据挖掘的价值在于以智能化方法发现数据背后的深层次规律,发掘现象之间的联系,如嫌疑人的兴趣爱好、行为偏好等。数据画像是传统犯罪心理画像在大数据时代的新发展,通过借助基础数据库及数据挖掘技术,大数据可以对嫌疑人进行全方位、多维度的数据刻画。犯罪网络关系分析主要应用于恐怖活动犯罪、毒品犯罪等有组织的犯罪,意在通过大数据技术来发现犯罪组织成员之间的关系及其分工合作情况。犯罪热点是分析犯罪活动在时空位置上的分布规律,大部分的犯罪往往集中在少部分地区;犯罪热点分析还往往与犯罪预测联系在一起,通过对历史犯罪热点数据的分析来预测未来犯罪活动的趋势和走向。在大数据侦查中,不能忽视大数据公司的作用,大数据公司所拥有的海量用户数据是侦查中的重要数据来源,侦查机关要积极寻求与大数据公司的数据共享及技术合作。

最后,本书论述了大数据侦查的相关制度构建,既包括大数据本身的法律程序构建,也包括与之相关的配套制度建设。从权利角度看,大数据侦查难免会侵犯公民的个人信息权。目前,刑事侦查中的个人信息保护尚属于法律真空地带。侦查机关的数据收集、数据共享不可避免地会形成“大数据监控社会”,带来民众的心理恐慌;侦查中对个人数据的二次分析、深度挖掘更是对个人信息权的严重侵犯。因此,本书从审查批准、个人参与、比例原则等方面去寻求大数据侦查与个人信息保护之间的价值平衡。从程序角度来看,大数据侦查过程是不透明的,当事人不知道自己的哪些数据被收集、被分析,也不知道自己被采取侦查措施的数据依据。可见,大数据侦查对传统的正当程序带来一定影响,剥夺了当事人的知情权、辩护权等权利。因此,本书从通知解释、赋予异议权、数据记录等几个角度去规制大数据侦查的正当程序。另外,笔者还从数据共享、技术应用以及行业

规范的角度提出了大数据侦查相关配套制度。在数据共享方面,要打破不同地域、级别、部门之间的数据孤岛现象,达到侦查机关内部的数据共享以及侦查机关与社会数据库共享;在技术方面,要建立大数据侦查的技术体系和应用平台;在行业规范方面,大数据公司要加强对个人数据的分级、分类管理,规范公权力机关调取数据的行为,对大数据公司的数据管理和第三方的数据调取进行衔接性规制。

本书系 2016 年国家社科基金年度项目“大数据时代电子文件的证据规则与管理法制建设研究”(项目批准号:16BFX033)阶段性研究成果。