

检察技术信息化系列教材

总主编 张雪樵 执行主编 赵志刚

检察实务中的大数据

PROCURATORIAL
BIG DATA

马建刚 著

中国检察出版社

检察技术信息化系列教材

总主编 张雪樵 执行主编 赵志刚

检察实务中的大数据

PROCURATORIAL
BIG DATA

马建刚 著

中国检察出版社

图书在版编目 (CIP) 数据

检察实务中的大数据/马建刚著. —北京: 中国检察出版社,
2017. 9

ISBN 978 - 7 - 5102 - 1930 - 6

I. ①检… II. ①马… III. ①数据处理 - 应用 - 检察机关 -
工作 - 中国 IV. ①D926. 3 - 39

中国版本图书馆 CIP 数据核字 (2017) 第 178860 号

检察实务中的大数据

马建刚 著

出版发行: 中国检察出版社

社 址: 北京市石景山区香山南路 109 号 (100144)

网 址: 中国检察出版社 (www. zgjccbs. com)

编辑电话: (010) 86423704

发行电话: (010) 86423726 86423727 86423728

(010) 86423730 68650016

经 销: 新华书店

印 刷: 河北省三河市燕山印刷有限公司

开 本: A5

印 张: 11. 125

字 数: 306 千字

版 次: 2017 年 9 月第一版 2017 年 9 月第一次印刷

书 号: ISBN 978 - 7 - 5102 - 1930 - 6

定 价: 48. 00 元

检察版图书, 版权所有, 侵权必究
如遇图书印装质量问题本社负责调换

《检察技术与信息化系列教材》

编委会

总主编：张雪樵

执行主编：赵志刚

委员：（按姓氏笔画排序）

王桂强	冯新	刘力	刘宝旭
刘建伟	刘勇	刘品新	朱修阳
江一山	严厚甫	张向阳	张春霞
时磊	李运策	李学军	周颂东
辛生	姚郑	贺德银	钟福雄
秦仲学	聂宏	贾茂林	黄华
程剑峰	缪存孟		

总 序

2017年是全面深化改革的关键一年，是全面深化司法体制改革决战之年。在全面推进依法治国、全面深化改革的当下，在人民群众对公平正义的追求越来越迫切的今天，《检察技术与信息化系列教材》得以陆续出版，对于全面推进科技引领检察工作现代化，是一件很有意义的事情。

科技与法治，对当代中国而言，非常重要。实现中华民族伟大复兴是当前中华民族最伟大的梦想，习近平总书记指出“科技兴则民族兴，科技强则民族强”，科技越来越成为经济社会发展的主要力量。最近，习近平总书记又对司法体制改革作出重要指示，强调“把深化司法体制改革和现代科技应用结合起来，不断完善和发展中国特色社会主义司法制度”。检察技术信息工作的功能已非同日而语，随之，相应的教育培训、规范运用也日趋重要。《检察技术与信息化系列教材》不仅是研究交叉学科，更是探索科技与检察工作如何在实践中深度融合。科学有着对真理的求索，法治有着对公正的追求，这是科技与法治在价值上的契合；科技与法治都需要高度的客观理性，一个要符合自然规律，一个要符合社会规律，这是科技与法治在思维上的吻合。同时，科技与法治的差异也十分明显，联系是融合的基础，差异是碰撞出火花的前提。正是这种联

系与差异，使得科技进步能够引领法治的进步，使得科技化可以促进检察工作的现代化。

为此，在世纪之初，最高人民检察院提出科技强检战略，编制这套教材也旨在通过科技手段加强检察工作。近年来，科技工作在检察工作中的战略性、基础性地位日益突出，核心战斗力作用得到充分发挥。最高人民检察院连续编制两个五年的科技强检规划，连续三年召开电子检务工程推进会，将科技强检战略摆到了前所未有的高度。曹建明检察长在第十四次全国检察工作会议和刚刚结束的大检察官研讨班上多次提到科技强检相关工作，要求在规划中重视、在实施中重视、在运用中重视，可以说重视程度是空前的。科学技术在检察机关应用的深度和广度不断提高，现代科技逐步融入各项检察工作，检察人员运用科技的意识和能力逐步加强，科技手段在强化司法办案、深化检务公开、提升司法公信力中的作用凸显。随着科技的进步，尤其是社会信息化的深入发展，以数字化、网络化、智能化为主要特征的创新浪潮席卷全球，新一代科学技术正在深刻改变着经济社会的发展模式和人民群众的生产生活方式，同时也对司法实践产生巨大影响，对旧有的证据规则、办案模式、公开方式等方方面面产生了巨大的冲击。指纹技术从发现到用诸于司法实践间隔了上千年，遗传理论从诞生到基因鉴定作为诉讼证据间隔了上百年，但计算机从诞生到信息化成果应用于检察办公办案，只用了几十年，由此可见科技发展之迅猛。

科学技术的进步会带动检察工作的进步，然而科学技术的进步，并不能自动地转化为检察工作的进步。科技强检战略的实施，不仅仅需要科技基础，更需要发挥主观能

动性，需要全体检察工作者共同努力，实现科技与检察工作的融合。

《检察技术与信息化系列教材》在编制过程中格外重视科技与检察工作的结合，紧跟科技前沿发展，又深入探索思考科技在检察工作中的现实应用，这是这套教材的主旨所在、价值之处。整个系列可以分为三类：

一是偏重自然科学的研究，是整套教材的基础。检察技术与信息化发展的根本动力是科学技术的发展。这类教材站在了时代前沿，紧跟科技潮流，为新技术转化为司法成果作好准备，在科技应用于检察工作之前，将科学原理弄清楚、搞明白，指导实践，帮助读者在应用的各个环节，遵循自然科学的规律。

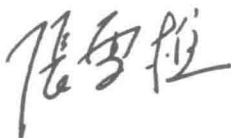
二是制度理论探索，是整套教材的核心。这类教材着重在科技应用层面上做功课、下工夫，强调科技与检察的深度融合，这也是检察技术与信息化工作的重点。当前，司法责任制改革、以审判为中心的诉讼制度改革，为科技应用提出了丰富复杂的应用场景和要求，相关配套措施的出台急需制度层面的设计，也需要科技应用层面的创新。我们如何应用、消化丰富的科技资源，为检察工作提供科技支撑，值得我们每一位检察工作者深思。

三是对实务操作规范加以摸索，是整套教材的关键。只有当个人经验超脱平面，建立起有效连接，在体系中积累，在碰撞中比较，在争鸣中发展，经实践反复检验才能逐渐形成共识，不仅实务类的教材具有这样的特点，整套教材都源于实践，臻于实践，尤其是有大量宝贵的基层实践经验。这类教材几乎涵盖了检察技术与信息化的各个门类的实践经验，无论是从事法医、文检、司法会计、心理

测试、同步录音录像，还是项目管理、视频系统、运维保障等，相信都能从中得到有益借鉴。

总的来说，《检察技术与信息化系列教材》，紧紧围绕科技强检这一主题，紧密联系检察工作实际，既有实务又有理论，既接地气又有高度，相信这一系列教材一定能发挥出指导学习方向，交流先进经验，探索重大理论的作用，与各位读者共同促进科技强检工作的进一步发展。丛书虽为教材，但一定还有纰漏之处需要完善，所幸教学相长，希望每位读者能够不吝意见，与编委会多多交流，同时也祝愿每位读者能够与《检察技术与信息化系列教材》一道，共同进步，共同成长。

最后我个人有一点期许，我希望能够借《检察技术与信息化系列教材》出版这一契机，不仅仅使得科技工作与检察工作深度融合，更使得科技思维与法治思维能够相得益彰，使得这种复合思维不仅在检察工作、司法工作中得以应用，也在公众之间广泛传播。



2017年9月1日

序

随着社会信息化的深入发展，新一代信息技术尤其是大数据正在深刻改变着经济社会的发展模式和生产生活方式。党的十八届五中全会对实施国家大数据战略作了全面部署。《国民经济和社会发展的第十三个五年规划纲要》提出了“推进国家治理体系和治理能力现代化”的总目标。大国治理需要大智慧，无数据不智慧，大智慧往往源于大数据。大数据作为信息时代的基础资源，为国家治理现代化提供基础数据和决策支撑。大数据的发展与应用，将对社会的组织结构、国家的治理模式等产生深刻影响，同时也深刻改变了信息化发展的技术环境条件。这对检察机关提高司法办案能力，提升司法为民的服务能力是一个挑战，同时也提供了难得的机遇。

最高人民检察院在《“十三五”时期科技强检规划纲要》中提出构建智慧检务体系，就是要在检察工作中加强应用大数据技术力度，促进检务智能化应用，实现精细化管理、科学决策和智慧服务，推进现代科技与检察工作深度融合，促进检察工作提质增效，深化检务公开，从而更好地规范司法行为、提升司法能力、强化法律监督、维护司法公正。

建刚是我指导的博士后，他在做学问上一直是踏实努力的。记得在中国人民大学博士后进站面试时，提及他从

中国科学院软件所计算机博士毕业跨专业来做法学博士后，需要加强法律方面的学习，他当即表示要考司法考试，半年后他兴奋地告诉我顺利通过了司法考试。2016年5月他又以博士后研究内容为主申报成功了国家博士后科学基金。6月中国人民大学博士后集中授课时，他说想把博士后在站期间的研究成果总结出来写成一本书，并拿了《检察实务中的大数据》的写作提纲给我看，我鼓励他抓紧完成。时至年底，他完成了书稿提交审阅。

目前，检察机关正在积极落实国家大数据战略，推动建设智慧检务体系，但是这方面系统性的研究还比较少，本书算是这方面的探索。本书从政策对策、技术理论和业务实践等角度来阐述大数据与智慧检务，探讨了大数据开放共享、隐私保护、大数据分析、知识库和“互联网+”等热点问题，给出了大量国内外司法系统应用大数据的典型案列，并提出了我国检察机关的对策建议。

大数据时代既是充满无限生机的时代，也是一切都有可能的时代，希望广大检察干警在工作中不断强化对现代科学技术的运用，增强运用信息化手段的行动自觉，养成运用大数据分析、解决问题的思维习惯，努力提高检察工作智能化水平。

姜新功
2016.12.22

目 录

基础概念篇

第一章 智慧检务	3
第一节 国内外的“智慧政务”和“智慧法院”建设	4
第二节 什么是“智慧检务”	5
第三节 智慧检务的总体框架	9
第四节 智慧检务的应对策略及展望	12
第二章 大数据是什么	15
第一节 为什么需要大数据技术	15
第二节 大数据概念的内涵	17
第三节 大数据的典型特征及价值链	19
第四节 从理论、技术和实践角度认知大数据	22
第五节 大数据在刑事司法领域的应用案例	25

政策对策篇

第三章 大数据的国内外政策分析和检察机关对策	31
第一节 国外的大数据政策和典型应用	31
第二节 国内的大数据政策和地方政府的探索实践	34
第三节 检察机关的大数据对策	41
第四章 大数据开放共享与检察机关对策	45
第一节 国外的数据开放共享现状	45

第二节	国内的大数据开放共享现状	50
第三节	检察机关的数据开放与对策	55
第五章	大数据隐私保护与检察机关对策	58
第一节	大数据隐私问题与类别	58
第二节	大数据隐私保护技术	61
第三节	隐私保护法律法规和行业规范	63
第四节	检察机关对公民大数据隐私的保护对策	71
第六章	“互联网+”检察工作	73
第一节	“互联网+”概念的内涵	73
第二节	关于“互联网+”的国家政策	74
第三节	国内外通过移动互联网开展政务服务的情况	76
第四节	检察机关探索开展“互联网+检察工作”的情况	78
第五节	检察机关开展“互联网+检察工作”的对策及展望	81
第七章	电子政务视野下的云计算国内外政策和检察机关对策	84
第一节	国外发达国家的云计算发展战略和特征	84
第二节	国内的云计算发展政策	87
第三节	云计算标准	90
第四节	国内地方政府云计算的发展探索	91
第五节	检察机关的云计算发展对策	92

技术理论篇

第八章	大数据处理与可视化技术	97
第一节	大数据处理	97
第二节	大数据可视化分析	107

第九章 人工智能	117
第一节 人工智能和人工神经网络	117
第二节 国内外人工智能国家政策	119
第三节 机器学习框架和分类	123
第四节 深度学习	125
第五节 典型应用案例	131
第十章 数据交换分发共享集成技术	134
第一节 面向大规模数据交换分发的发布/订阅系统 技术	134
第二节 数据抽取转换加载工具 (ETL)	163
第三节 基于发布/订阅方式的语义数据交换分发集成 系统	165
第十一章 数据知识化相关技术	204
第一节 知识表示技术	204
第二节 基于知识的推理技术	215
第三节 知识可视化技术	216
第十二章 云计算技术	218
第一节 云计算概念的内涵	218
第二节 云计算的特征	220
第三节 云计算体系架构	221
第四节 虚拟化技术与存储技术	227
第五节 云计算安全	228

业务实践篇

第十三章 大数据与信息引导的职务犯罪侦查	233
第一节 基于大数据的职务犯罪侦查的研究背景与意义	234
第二节 国内外的研究和应用现状	236

第三节	大数据环境下职务犯罪向智慧侦查转型发展的 展望	245
第十四章	“互联网+”时代背景下的检务公开	250
第一节	国家政策要求检务公开顺应“互联网+”时代的 发展	250
第二节	互联网时代国内外网上司法公开的现状	251
第三节	检察机关网上检务公开系统建设应用中存在的 问题	255
第四节	“互联网+检务公开”的内涵、特征及系统建 设应用需注意的问题	256
第五节	检务公开系统的评价指标	257
第六节	检务公开系统的管理机制	260
第十五章	大数据环境下的涉检网络舆情监测系统	262
第一节	涉检网络舆情的概念和分类	263
第二节	网络舆情监测系统的总体架构	264
第三节	网络舆情的动态信息采集	265
第四节	涉检网络舆情信息的分布式存储	268
第五节	涉检网络舆情的数据处理	269
第六节	基于语义网的涉检网络舆情语义关联模型构建	272
第七节	网络舆情信息的查询与展示	273
第十六章	面向智慧检务的知识库构建	275
第一节	对电子卷宗库采用 NLP 方法抽取新知识	275
第二节	基于统一业务应用系统进行关系数据库知识 抽取	278
第三节	检察行业信息标准化	280
第四节	多源知识融合和基于知识库的查询、统计及 预测	282

第十七章	检察机关与公安、法院信息共享业务协同平台·····	285
第一节	公检法信息共享的研究应用背景·····	285
第二节	公检法业务交互模型·····	286
第三节	系统框架体系结构·····	287
第四节	安全支撑平台及信息共享标准规范·····	292
第五节	方案特点分析和实施建设需要注意的问题·····	293
第六节	典型应用案例·····	295
第十八章	检察机关行政执法与刑事司法衔接信息共享平台·····	296
第一节	“两法衔接”信息共享平台的研究应用背景及业务交互模型·····	296
第二节	系统框架体系结构·····	298
第三节	系统的主要设计·····	300
第四节	方案的实施效果分析及实施建设需要注意的问题·····	307
第十九章	大数据与人工智能在推进智慧检务体系建设中的应用展望·····	309
第一节	智慧检务建设的时间表和路线图·····	309
第二节	利用大数据与人工智能技术开展智慧检务应用的基础·····	310
第三节	大数据与人工智能技术在各项检察业务中的应用展望·····	311
第四节	总结展望·····	316
参考文献	·····	318

基础概念篇

