

互联网+大数据 在执法办案中的应用

网络使用篇 网络爬虫篇 电子证据篇 取证技术篇

孙百昌◎著

 中国工商出版社

互联网+大数据 在执法办案中的应用

网络使用篇 网络爬虫篇 电子证据篇 取证技术篇

孙百昌◎著

 中国工商出版社

责任编辑 / 刘安伟 聂艳

封面设计 / 木籽

图书在版编目 (CIP) 数据

互联网 + 大数据在执法办案中的应用 / 孙百昌著. —
北京 : 中国工商出版社, 2016.6

ISBN 978-7-80215-873-3

I . ①互… II . ①孙… III . ①互联网络 - 应用 - 行政
执法 - 研究 - 中国 IV . ① D922.114-39

中国版本图书馆 CIP 数据核字 (2016) 第 147806 号

书名 / 互联网 + 大数据在执法办案中的应用

著者 / 孙百昌

出版 · 发行 / 中国工商出版社

经销 / 新华书店

印刷 / 北京翌新工商印制公司

开本 / 787 × 1092 毫米 1/16 印张 / 21.375 字数 / 270 千字

版本 / 2016 年 8 月第 1 版 2017 年 5 月第 2 次印刷

社址 / 北京市丰台区花乡育芳园东里 23 号 (100070)

电话 / (010) 63730074, 63725178 电子邮箱 / zggsccbs@163.com

出版声明 / 版权所有, 侵权必究

书号: ISBN 978-7-80215-873-3/D · 519

定价: 58.00 元

(如有缺页或倒装, 本社负责退换)

本书着眼于执法办案的实用性。一些增强理解的理论要点、解释说明等放在了注释上，供有兴趣的读者参考。

笔者《大数据时代市场监督管理——分析方法与范例》2015年9月出版后，互联网+大数据在市场监管中的运用进一步快速发展，可以用“风起云涌”来形容。¹办案方面使用互联网+大数据的呼声尤其高。在这种形势下，办案模式的转变已初见端倪。工商、质监、食药监等市场监管部门开始探讨、实践从过去习惯依靠传统手段办案向尝试运用现代信息技术转变。在需求和政策的双重推动下，互联网和大数据与办案之间的脉络逐渐凸显出来。这种状态，给笔者提供了丰富的资料，也给了接地气的启发。

为什么一定要把互联网和大数据用于执法办案呢？简单地说，时代确实不同了！具体说有两点：第一，市场主体数量剧增；第二，市场交易模式大变。

在市场监管中，被监管的对象越多，需要的人力、物力越多，到一定程度监管能力会遇到瓶颈。与此同时，现代市场交易模式通过互联网+大数据已经有了极大延伸。

在市场主体数量剧增、市场交易模式大变的情况下，如何在中国进行好的市场监管，互联网+大数据提供的零边际成本条件为监管模式提

1 “风起”是指在国家层面的推动下，担负市场监管的工商、质监、食药监和合并后的市场监管总局，以及各个基层监管部门，纷纷表现出运用互联网和大数据实施市场监管的极大兴趣。各种典型案例、做法、探讨陆续出现；“云涌”是指大数据的发展极其快速，好似暴风雨前的云卷浪翻。互联网和大数据越来越“火”了，中央领导不断提互联网+大数据，国务院文件一个接一个公布，各个部门陆续出台了运用大数据于政府管理的意见、规划等。上行下效，红红火火。

升提供了机遇，²互联网快速发展的今天，我们的社会在接近零边际成本社会。在零边际成本社会中，市场监管部门可以也可能使用互联网+大数据以接近免费的方式获取市场主体及其行为的信息，包括执法办案需要的信息，这是最具有生态效益的市场监管和执法办案新模式，也是符合时代特点的可持续发展的模式。传统监管模式的非数字化实体监管必须也有可能延伸，升级为数字化监管。正因为如此，高速互联网下的大数据运用已经是发达国家市场管理的基础和基本资源。加拿大学者麦克卢汉认为，人类发展就是人通过技术手段不断延伸自己的肢体器官：望远镜是眼的延伸、收音机是耳的延伸、广播是嘴的延伸、机器人是手的延伸、飞机是腿的延伸、计算机是大脑的延伸……互联网给市场监管带来巨大的变革，它的实质在于延伸了监管者的眼、耳、嘴、手、腿，甚至替代了部分判断。科学技术是第一生产力，当监管和互联网+大数据结合后，监管力会有极大的提高。随着互联网+大数据的运用，体制、机制整合成为刚需，成为驱动市场监管改革的动力。这就是为什么国务院要求要充分认识运用大数据加强我国市场主体服务和监管的重要性，迎接大数据对政府服务和监管能力提出的新挑战，抓住这个新机遇。

从办案的角度说，和我们以往办案的方法相比，互联网+大数据用于办案，就像在市场大海中捕鱼，而现在我们办案的方法，就像在池塘里捕鱼。现在的市场，已经是全球化的大海，画地为界的池塘市场已经成为历史。利用互联网和大数据办案的“大海捕鱼”比喻，至少可以体会出三点：第一，范围大小不同。大海无边无际，一眼望不到边；池塘范围有限，都在捕鱼人的视野中。这意味着在互联网中利用大数据办案，先要放开视野，如果视野放不开，是不能体会到大海捕鱼乐趣的。第二，数量多少不同。

2 边际成本指的是每一新增单位引起的成本。Jeremy Rifkin 在其著作《零边际成本社会：一个物联网、合作共赢的新经济时代》中提出了“零边际成本社会”的概念。他指出，在未来时代，社会协同与科技进步将共同打造出“零边际成本社会”，这将更多地改变社会组织运行方式，人类社会的市场机制会有更多的共享成分。在零边际成本社会中，传统的市场监管模式几乎完全失效，这是因为传统模式是对市场个体点对点的监管，而发达的市场协作共享将导致点对点模式不可能实施有效地全面监管。一个明显的证据就是市场监管中的食品监管的失灵。

大海中的鱼量几乎无限，近海捕不着了去远海，这种鱼没了捕其他鱼。池塘的鱼有限，不用捕也知道有多少。这就像现在的办案，基层大致都知道这个地方有多少案子。所以经常会为案源费神。利用大数据办案，案源几乎无限，网络延伸到哪里，案源就可以延伸到哪里。更重要的是现在许多市场行为都发生在网络上，新型案源多在网上。第三，办案方法不同。利用互联网和大数据办案需要一定的技术能力，知识在这里体现了巨大的价值，科技成为市场监管的重要因素。沿着这条路走下去，执法办案就会走上符合现代社会发展趋势的道路。

互联网和大数据就像大海，办案就像打渔。这样有四个问题：第一个问题是如何使用大海资源；第二个问题是如何收集大海资源；第三个问题是如何打鱼和如何处理打来的鱼；第四个问题是研究打渔的工具，“工欲善其事，必先利其器”。这四个问题，形成了本书的四篇：网络使用篇、网络爬虫篇、电子证据篇、取证技术篇。最后的附录列举了案例、对办案的作用等索引。

本书适用于工商行政、质量监督、食药监督、物价等市场监管部门利用互联网+大数据执法办案。

阅读本书可解决以下办案中的问题：（1）指引执法者充分利用互联网资源发现案源、理解法律、提供办案参照；（2）简单而快速地学会使用网络爬虫进行网络巡查和自动发现违法线索；（3）快速学会电子证据的发现、提取和固定的方法；（4）有兴趣的可以进一步学习掌握计算机取证的软硬件基础知识。

孙百昌

二〇一六年五月

目 录

引 言

第一章 执法办案面临的挑战

1.1 传统执法办案模式	2
1.1.1 传统执法办案模式的三个特点	2
1.1.2 简单分析一下原因	4
1.2 原有靠人力和经验的执法办案模式走不下去了	6
1.2.1 监管比严重失衡	6
1.2.2 市场形态的变化	8
1.2.3 监管效率在下滑	12
1.3 大数据的主要特征	15
1.3.1 大量	16
1.3.2 多样	17
1.3.3 价值	17

网络使用篇

第二章 行政处罚如何借鉴互联网和大数据

2.1 中国裁判文书网	21
2.1.1 两种检索方法	22
2.1.2 关键字的选择	24
2.2 执法主体问题	25
2.2.1 法律没有授权行政部门最终裁量权	25

2.2.2 部门管辖争议	26
2.2.3 地域管辖争议	28
2.2.4 级别管辖争议	30
2.2.5 移送管辖参考	32
2.3 法律理解问题	34
2.4 法律适用问题	38
2.5 执法程序问题	40
2.6 执法责任问题	42
2.6.1 玩忽职守罪	43
2.6.2 滥用职权罪	44
2.6.3 徇私舞弊不移交刑事案件罪	45
2.6.4 食品安全监管渎职罪	45

第三章 利用食品药品信息查办案件

3.1 可利用的网络资源	48
3.2 核对食品生产许可证的方法	49
3.2.1 核对 SC 证的方法	49
3.2.2 核对 QS 证的方法	49
3.2.3 伪造食品生产许可证编号案件	51
3.2.4 冒用食品生产许可证编号案件	51
3.3 使用过期食品生产许可证编号案件	52
3.4 查询食品添加剂生产许可获证企业	52
3.5 查询保健食品	53
3.6 查询虚假广告	54
3.7 药品、医疗器械、化妆品查询	55
3.8 互联网药品信息服务、交易服务和网上药店案件	56
3.8.1 两种服务和 A、B、C 证书	56
3.8.2 检查方法	58
3.9 食药监总局“食品安全国家标准数据检索平台”	61
3.10 利用条形码和国家食品安全追溯平台查询	63
3.11 婴幼儿配方乳粉追溯平台	64

第四章 利用质量标准查办案件

4.1 利用强制性国家标准查办案件	68
4.1.1 可以利用的网站	69
4.1.2 玩具标准	70
4.1.3 玩具的标签标准	70
4.1.4 能不能目测认定	71
4.1.5 确定法律责任的依据	73
4.1.6 归纳一下思维逻辑	74
4.2 利用推荐性国家标准查办案件	75
4.2.1 查找推荐性国家标准	76
4.2.2 标准关键原理	77
4.2.3 标准使用步骤	77
4.2.4 抽样结果判定	79
4.3 利用行业标准查办案件	80
4.4 条形码和食品产品信息	82

第五章 利用认证认可信息查办案件

5.1 认证认可概述	86
5.2 可以利用的网络资源	87
5.3 擅自从事认证活动	88
5.4 超出批准范围从事认证活动	89
5.5 聘用未经认可机构注册的人员从事认证活动	90
5.6 查询产品强制性认证证书是否真实	91
5.7 列入目录的产品未经认证擅自出厂或销售	92

第六章 利用产品召回信息查办案件

6.1 涉及召回的法律规范	96
6.1.1 对特定产品召回制定的法律规范	96
6.1.2 一般性召回法律规范	97

6.2 查询召回信息	98
6.2.1 利用国家质检总局缺陷产品管理中心网站	99
6.2.2 利用国家食药监总局和各地食药监局网站	101
6.3 缺陷汽车召回案件	102
6.3.1 重要性	102
6.3.2 查找召回信息	103
6.3.3 法律责任	104
6.4 药品召回案件	104
6.4.1 基本概念	105
6.4.2 查找召回信息	105
6.4.3 法律责任	106
6.5 医疗器械召回案件	108
6.5.1 基本概念	108
6.5.2 查找召回信息	108
6.5.3 法律责任	109
6.6 缺陷消费品召回案件	110
6.6.1 查找召回信息	111
6.6.2 法律责任	112
6.6.3 召回目录	113
6.7 儿童玩具召回案件	113
6.7.1 查找召回信息	114
6.7.2 法律责任	114
6.8 食品召回案件	116
6.8.1 食品召回的特殊性	116
6.8.2 查找不合格食品信息	117

网络爬虫篇

第七章 使用网络爬虫搜集执法办案所需信息

7.1 网络爬虫和其特点	120
7.2 直观使用网络爬虫的一个例子	122

7.2.1 八爪鱼采集器的简单使用	122
7.2.2 自动采集信息	126
7.3 自行编写爬虫程序	127

第八章 网络爬虫的使用：文本单页抓取

8.1 下载和安装	131
8.2 基本功能介绍	133
8.3 建立一个简单网页文本抓取任务	134
8.4 第一步，登录	135
8.5 第二步，新建任务	135
8.5.1 火车采集器主窗口	136
8.5.2 新建任务规则对话框	136
8.6 第三步，获取网址	137
8.6.1 获取单个网址	138
8.6.2 获取多个网址	138
8.6.3 获取有变化规律的网址	138
8.6.4 一个实例	139
8.7 第四步，设定单页网址采集规则	139
8.8 第五步，设定内容采集规则	140
8.8.1 内容采集规则选项卡介绍	140
8.8.2 什么是网页源代码	144
8.8.3 选择开头和结尾字符串	145
8.8.4 极端用语的智能筛选	147
8.9 第六步，设定内容发布规则	148
8.10 第七步，运行任务	151
8.11 抓取内容的使用	152

第九章 网络爬虫的使用：文本分页抓取

9.1 第一步，登录	155
9.2 第二步，新建任务	155

9.3 第三步，获取网址：多页网址的规律性发现	156
9.4 第四步，设定分页网址采集规则	157
9.5 第五步，设定内容采集规则	160
9.6 第六步，设定内容发布规则	164
9.7 第七步，运行任务	167
9.8 抓取内容的使用	167

第十章 网络爬虫的使用：图片抓取

10.1 第一步，登录.....	170
10.2 第二步，新建分组和新建任务.....	170
10.3 第三步，获取图片网址.....	172
10.4 第四步，设定网址采集规则.....	173
10.5 第五步，设定内容采集规则.....	174
10.6 第六步，设定内容发布规则.....	176
10.7 第七步，其他设置.....	178
10.8 第八步，运行任务.....	178

电子证据篇

第十一章 电子数据作为证据

11.1 电子证据及其概念.....	183
11.2 电子证据的法律概念.....	184
11.3 法律依据对使用电子证据的影响.....	185
11.4 如何识别电子证据.....	186
11.5 电子证据存放载体.....	187
11.6 电子证据的分布范围.....	188
11.7 电子证据转化为传统证据.....	189
11.8 应对电子证据质证.....	190

第十二章 电子证据的特征和取证标准

12.1 电子证据的 3 个特征	194
12.1.1 形态不直观	194
12.1.2 内容可改变	195
12.1.3 文件可删除	195
12.2 电子证据取证的 3 点注意事项	195
12.2.1 用多种方式取证	195
12.2.2 用电子签名固定证据	196
12.2.3 要搜索隐藏的文件	197
12.3 电子证据取证原则和分析要求	197
12.3.1 电子证据的取证原则	197
12.3.2 电子证据正本和副本的一致性检验	198
12.4 电子证据取证的行业标准	199
12.5 电子证据取证的其他标准	201

第十三章 现场电子证据取证

13.1 简易方法和完备方法	204
13.1.1 简易方法	204
13.1.2 完备方法	205
13.2 制作现场笔录	205
13.3 用照片记录电子证据取证现场	206
13.4 现场封存电子证据载体	207
13.5 现场获取动态电子证据	208
13.5.1 已经打开的文档	208
13.5.2 剪切板内容	208
13.5.3 聊天记录	210
13.5.4 浏览器中的收藏夹和历史记录	210
13.5.5 远程桌面配置	211
13.5.6 正在运行的云盘	212

13.5.7 Word 临时文件	212
13.5.8 内存中的数据	213
13.5.9 记录网络数据	213
13.6 屏幕录像	214
13.6.1 使用 Windows 操作系统自带的步骤记录器	215
13.6.2 使用 Mac（苹果电脑）自带的 QuickTime 进行屏幕录像	216
13.6.3 使用其他屏幕录像软件进行屏幕录像	217
13.7 复制电子证据	219
13.8 补强证据、确定当事人、询问笔录	220

第十四章 网络电子证据取证

14.1 什么是网络电子证据	223
14.1.1 网络电子证据概念	223
14.1.2 网络电子证据取证	223
14.1.3 网络电子证据特点	223
14.2 如何把看到的网页作为证据留存下来	224
14.2.1 截屏和保存	224
14.2.2 录像和保存	226
14.3 如何确认网页的所有者（当事人）	227
14.3.1 基本思路	227
14.3.2 域名和 IP 地址	227
14.3.3 通过网页查找当事人的步骤	230
14.3.4 通过网卡和 MAC 找到具体当事人	232
14.4 如何识别作为证据留存的网页没有被改动	232
14.5 关于网络的延伸知识	233
14.5.1 网络基本结构	233
14.5.2 TCP/IP 协议	233

第十五章 其他电子证据取证

15.1 手机电子证据取证	236
---------------------	-----

15.1.1 手机电子证据的存储介质	236
15.1.2 取证设备和工具软件	237
15.1.3 手机电子证据的提取	237
15.2 电子邮件证据取证	238
15.2.1 电子邮件概述	238
15.2.2 电子邮件取证	238

取证技术篇

第十六章 电子证据取证的硬件和软件基础

16.1 计算机的主要硬件	254
16.1.1 主板	254
16.1.2 CPU	255
16.1.3 内存	255
16.1.4 硬盘	256
16.1.5 常见的硬盘接口类型	258
16.1.6 多磁盘卷	259
16.2 计算机中的信息编码	259
16.2.1 二进制	259
16.2.2 其他进制	260
16.2.3 字符编码	260
16.2.4 编辑工具	261
16.2.5 文件乱码的处理	262
16.3 硬盘的三种分区	262
16.3.1 硬盘的三种分区	262
16.3.2 检查和显示隐藏的分區	264
16.4 硬盘文件系统结构	264
16.4.1 主引导扇区 (MBR、DPT)	265
16.4.2 操作系统引导扇区 (OBR、BPB)	266
16.4.3 文件分配表 (FAT)	266

16.4.4 目录区 (DIR)	267
16.4.5 数据区 (DATA)	268
16.5 硬盘格式化和数据恢复	268
16.5.1 硬盘低级格式化和数据恢复	269
16.5.2 高级格式化	269
16.5.3 数据恢复	270
16.6 BIOS	271
16.6.1 进入 BIOS 设置	271
16.6.2 Advanced BIOS Features 子菜单	273
16.6.3 Integrated Peripherals 子菜单	273
16.7 计算机操作系统	274
16.7.1 主流操作系统	274
16.7.2 Windows 系统启动过程	274
16.7.3 Windows 操作系统注册表	275
16.7.4 Windows 操作系统的日志文件	276
16.8 数据库基础	279
16.8.1 数据库的结构	279
16.8.2 查询数据库的 SQL 语言	280
16.8.3 主流数据库类型	280

第十七章 电子证据取证技术

17.1 数据删除和简要恢复原理	282
17.1.1 数据存取删除原理	282
17.1.2 数据恢复的基本操作步骤	284
17.1.3 Windows 系统数据删除和恢复	284
17.1.4 UNIX 系统数据删除和恢复	285
17.1.5 数据恢复的前提是数据不能被覆盖	285
17.1.6 数据恢复的理论基础是数据结构	287
17.2 数据恢复软件	287
17.2.1 EasyRecovery	288

17.2.2 WinHex	288
17.2.3 FinalData	288
17.2.4 R-Studio	288
17.2.5 RecoverNT	289
17.2.6 DiskForen	289
17.3 常见的简单加密方法	289
17.3.1 加密开机：设置 BIOS 密码	289
17.3.2 加密操作系统：设置进入 Windows 密码	290
17.3.3 利用软件加密	290
17.3.4 最简单的文件加密方法：修改后缀名	291
17.4 简单加密的破解方法	292
17.4.1 破解开机加密	292
17.4.2 破解操作系统进入密码	294
17.4.3 破解利用软件加密的文件或文件夹	295
17.5 哈希校验	296
17.5.1 电子证据的原始性校验	296
17.5.2 对创建了哈希值的电子证据的处理	298
17.5.3 获取哈希校验软件	298
17.6 克隆硬盘取证	299
17.6.1 用软件克隆整个硬盘	299
17.6.2 用软件克隆硬盘分区	301
17.6.3 使用硬盘克隆机	303
17.6.4 计算硬盘的哈希值	303
17.7 检查 Cookie 的软件	305
17.8 电子证据综合取证软件	306
17.8.1 开源取证工具 Autopsy	306
17.8.2 国内电子证据取证软件	306
17.8.3 Encase	307
17.8.4 FTK (ForensicToolkit)	307
17.8.5 X-Ways Forensics	308
17.8.6 WinHex	308