



大数据时代

数据保护与流动规则

王融◎著

THE RULES OF DATA PROTECTION
AND FLOW IN THE DIGITAL ERA

人民邮电出版社

北京

图书在版编目 (C I P) 数据

大数据时代：数据保护与流动规则 / 王融著. --
北京：人民邮电出版社，2017.3
ISBN 978-7-115-43919-2

I. ①大… II. ①王… III. ①数据保护 IV.
①TP309.2

中国版本图书馆CIP数据核字(2017)第044456号

内 容 提 要

本书立足实务，聚焦前沿问题，系统介绍了世界主要国家及地区个人数据保护法律规则、执法体系，并通过近40个鲜活案例生动展现了个人数据保护法律规则在实践中的具体运用。围绕大数据背景下的数据保护困境，如有效的同意机制、数据匿名化、数据跨境流动等现实问题，全面阐述法律规则最新进展，为企业开展大数据商业实践提供可具操作性的合规指南，也为政府部门推进相关领域制度建设提供参考。

◆ 著 王 融

责任编辑 吴娜达

执行编辑 唐名威

责任印制 彭志环

◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路11号

邮编 100164 电子邮件 315@ptpress.com.cn

网址 <http://www.ptpress.com.cn>

固安县铭成印刷有限公司印刷

◆ 开本：700×1000 1/16

印张：19

2017年3月第1版

字数：311千字

2017年3月河北第1次印刷

定价：89.00元

读者服务热线：(010)81055488 印装质量热线：(010)81055316
反盗版热线：(010)81055315

Preface

序

人类社会的发展先后经历了农业时代、工业时代和我们正身处其中的信息时代。每一个时代，人类都创造了支撑经济社会有效运转的制度规则。农业时代，围绕土地建立了以不动产、动产为基本对象的物权规则，使耕者有其田，居者有其屋。英国工业革命的发生，一个重要的制度基础在于确立了知识产权保护法律制度，由此调动了发明人的积极性，直接推动了蒸汽机等技术发明创造活动；有限责任公司制度法律制度的确立，使创新、创业活动免于后顾之忧，企业家精神得以走上历史舞台。进入信息时代，作为生产要素重要组成部分的信息资源不同于传统的生产要素，不再具有排他与独占等特性，由此使信息的生成、收集、使用规则成为新的制度核心。如何建立科学合理的个人信息保护规则，以在保证个人尊严与自由的前提下，实现对信息的充分流转和使用、促进经济增长和社会进步，成为人们共同的追问。幸运的是，我们这一代人正在见证乃至深度参与这些规则建设。

早在十年以前，我就开始关注世界各国的个人信息保护法律制度，并将其作为自己的重点研究领域之一。随着研究的深入，我越来越清楚地认识到，越是经济及技术发达的国家，越是重视个人信息保护问题，其法律制度也愈加完备。

经过几十年来、尤其是最近几年来的努力，我国的个人信息保护法律制度逐渐完善，国内关于个人信息保护的研究也日渐深入。在立法实践方面，以《中华人民共和国网络安全法》《中华人民共和国侵权责任法》《全国人大常委会关于加强网络信息保护的決定》《中华人民共和国刑法修正案（九）》等法律为框架，我国已经初步建立了涵盖民事救济、行政监管、刑事追责的个人信息保护法律体系；在法律研究方面，个人信息保护已经成为学术研究

的重点领域，国内关于个人信息保护的研究成果也日渐丰厚。更加令人欣喜的是，个人信息保护问题也日渐成为全社会广泛关注的问题。有关个人信息保护的事件、判例、政策进展成为人们关注热点，也是每年两会人大代表、政协委员献言献策的主要议题。在学术界、产业界吸引了一大批学者、从业者投身于该领域的研究，从法律、技术、经济、社会、人文等多个视角推动对个人信息保护问题的不断深化。

王融是近年来在个人信息保护领域持续研究、不断积累的一位学术新星。她基础扎实、知识全面。尤其因为工作关系，保持了对互联网技术、业务、产业最密切的跟踪和了解，对移动互联网、云计算、大数据等新技术新业务带来的个人信息保护制度适用问题有着细致的分析和独到的洞察，因此也往往能够与实务结合得更加紧密。

《大数据时代 数据保护与流动规则》是王融近年来在此领域研究工作的积累，它不仅涵盖了对全球个人信息保护法律规则的详细介绍，更是纳入了各国及地区的典型执法、司法案例，为理解一国/地区个人信息保护法规政策提供了鲜活的注解，能够对实务工作起到更加直接的参考作用。

本书的另外一个亮点是重点问题突出，对个人信息保护政策在大数据时代的新问题、新对策、新进展系统阐述，并对当下最关注的个人信息的界定、同意机制、数据匿名化、数据权属、数据本地化问题进行了有价值的探讨。

期待这本书能够为管理部门、企事业单位、高校研究机构工作的相关人士，在制定、实施、研究个人信息保护法律政策工作中提供积极帮助。

是为序。

周汉华*

2017年3月于北京

* 法律学者，中国社会科学院法学研究所研究员。

Introduction

前言

大数据时代，数据的收集、分析与利用构成企业运营的核心。在这一历史浪潮中，除了技术进步和商业创新为大数据发展提供基本支撑和原生动力外，法律规则也将扮演重要角色。如何合法合规地收集、利用数据是大数据产业链上每一个主体所面临的首要问题。

当前，数据保护规则正在全球加速建立。截至 2015 年 2 月，已有 109 个国家或地区制定了个人数据保护法律和规章，超过了全球拥有司法管辖权地区一半以上，并且这一数字仍在持续增长。数据保护执法机构在全球普遍设立，执法力度与权威不断加强，成为大数据时代高悬在公司、政府机构乃至个人之上的达摩斯利剑。数据从收集到加工、从交易到披露、从境内到境外的转移流动，每一个处理环节都会在法律规则的约束之下。大数据推动了数据处理活动的几何级指数增长，同时也促使数据保护规则的不断完善与跟进。

与大数据的日新月异发展速度相比，我国对相关数据保护规则的了解则几乎处于起步阶段。消费者、企业、政府监管机构等各方主体对于数据规则的认识并不清晰，在有些领域，甚至存在着规则混淆甚至错误的情况。

正因如此，本书旨在系统介绍当前数据保护的基本规则，并通过典型案例的方式，深入浅出地反映数据保护法律和规章原理与具体规范，为企业开展大数据应用与商业实务提供具有可操作性的合规指南，为政府等监管机构了解境外数据规则体系、制定符合国情的数据保护规则提供重要参考。

本书共分为 5 章。第 1 章为大数据概述，简要介绍大数据诞生背景、应用发展以及其带来的隐私风险。第 2 章为个人数据保护规则概述，系统介绍数据保护规则与大数据之间的关系、数据保护法规的兴起、发展、核心概念与主要原则。第 3 章介绍全球典型国家、地区个人数据保护法规制定及执行

基本情况。第4章介绍大数据时代下数据保护新规则，结合大数据产业和应用发展最新态势以及各国或地区最新的法规执行实践，对新的法律和规章制度进行重点介绍，为企业如何做好数据保护合规工作提供建议。第5章讲述数据跨境流动规则，介绍国际上数据跨境流动基本规则和近年来的数据本地化政策，以及这些政策对企业合规带来的影响。

王 融

2016年5月于北京

Contents

目录

序 // I

前言 // III

第1章 大数据时代 //001

- 1.1 大数据诞生的背景 //001
- 1.2 大数据的含义 //002
- 1.3 大数据应用发展 //002
- 1.4 大数据带来隐私风险 //005

第2章 个人数据保护规则概述 //011

- 2.1 大数据与数据保护规则 //011
 - 2.1.1 概述 //011
 - 2.1.2 二者之关系 //012
- 2.2 个人数据保护法规概述 //014
 - 2.2.1 个人数据保护法规制定的兴起 //014
 - 2.2.2 个人数据保护法规制定的两种基本模式：差异与妥协 //016
 - 2.2.3 个人数据概念 //018
 - 2.2.4 个人数据保护的主体 //027
 - 2.2.5 个人数据保护法规的主要原则 //029

第3章 全球典型国家、地区个人数据保护法规制定及执行情况 //035

- 3.1 美国 //035
 - 3.1.1 主要立法 //036

- 3.1.2 主要执法机构——美国联邦贸易委员会 (FTC) //041
- 3.1.3 典型案例 //043
- 3.1.4 其他行业的执法机构——以电信业为例 //047
- 3.1.5 典型案例 //049
- 3.2 欧盟 //050
 - 3.2.1 主要立法 //051
 - 3.2.2 主要执法机构 //056
 - 3.2.3 典型案例 //060
- 3.3 英国 //063
 - 3.3.1 主要立法 //063
 - 3.3.2 主管机构 //065
 - 3.3.3 典型案例 //066
- 3.4 韩国 //068
 - 3.4.1 主要立法 //068
 - 3.4.2 主管机构 //075
 - 3.4.3 典型案例 //076
- 3.5 日本 //080
 - 3.5.1 主要立法 //081
 - 3.5.2 主管机构 //087
 - 3.5.3 典型案例 //088
- 3.6 新加坡 //088
 - 3.6.1 主要立法 //089
 - 3.6.2 主管机构 //094
- 3.7 中国 //097
 - 3.7.1 主要立法 //099
 - 3.7.2 主管机构 //108
 - 3.7.3 典型案例 //108
- 3.8 中国香港 //114
 - 3.8.1 主要法规 //114
 - 3.8.2 主管机构 //119
 - 3.8.3 典型案例 //120
- 3.9 中国澳门 //122
 - 3.9.1 主要法规 //122
 - 3.9.2 主管机构 //126

3.9.3 典型案例 //128

3.10 中国台湾 //133

3.10.1 主要法规 //133

3.10.2 主管机构 //139

3.10.3 典型案例 //141

第4章 大数据时代，个人数据保护新规则 //147

4.1 背景：悄然变化的网络世界与隐私观念 //147

4.2 个人数据保护法规的适用困境与问题 //150

4.3 应对大数据挑战，全球个人数据保护法规制定与政策应对 //158

4.3.1 概述 //158

4.3.2 欧盟立法与政策应对 //162

4.3.3 美国立法政策与应对 //181

4.4 大数据时代，个人数据保护规则的新变化 //189

4.4.1 个人数据保护法规适用的地域范围扩展 //189

4.4.2 数据保护监管机构增多，监管权威提升 //191

4.4.3 “数据保护官”逐步成为企业的法定标配 //192

4.4.4 引入“被遗忘权” //193

4.4.5 引入“数据可携权” //210

4.4.6 隐私保护设计制度 //211

4.4.7 数据泄露通知制度 //214

4.4.8 对敏感数据的特殊保护 //217

4.4.9 针对大数据分析、数据画像的特别规范 //225

4.4.10 数据匿名化 //226

4.5 大数据时代，公司个人数据保护规则合规要点 //230

4.5.1 数据保护合规工作的重要性 //230

4.5.2 数据保护合规的主要方面 //232

4.5.3 当前数据保护合规的主要困惑与对策 //235

第5章 数据跨境流动规则 //245

5.1 概述 //245

5.2 数据跨境流动典型政策 //247

5.2.1 联合国《电脑处理数据文件规范指南（1990）》 //247

5.2.2 OECD《保护个人信息跨国传送及隐私权指导纲领（1980）》 //247

- 5.2.3 《APEC 隐私保护框架》与跨境数据流动规则 (CBPR) //252
- 5.2.4 欧盟跨境数据流动政策 //267
- 5.2.5 典型案例 //278
- 5.3 近年来数据本地化政策趋势 //279
 - 5.3.1 概述 //279
 - 5.3.2 典型国家的数据本地化政策 //280
 - 5.3.3 典型案例 //283
- 5.4 对公司合规的影响 //285

后记: 我国《个人信息保护法》立法前路 //289

第1章 大数据时代

1.1 大数据诞生的背景

半个世纪以来，随着计算机技术全面融入社会生活，信息爆炸已经积累到了一个开始引发变革的程度。社会不仅充斥着比以往更丰富的信息，而且信息增长的速度也在加快^[1]。2000年前后，互联网网页爆发式增长，每天新增约700万个网页，到2000年底全球网页数达到40亿¹，用户检索信息越来越不方便。谷歌等公司率先建立了覆盖数十亿网页的索引库²，开始提供较为精确的搜索服务，大大提升了人们使用互联网的效率，这是大数据应用的起点。

当时搜索引擎要存储和处理的数据，不仅数量之大前所未有的，而且以非结构化数据为主，传统技术无法应对。为此，谷歌提出了一套以分布式为特征的全新技术体系，即后来陆续公开的分布式文件系统（Google File System, GFS）、分布式并行计算（MapReduce）和分布式数据库（BigTable）等技术，以较低的成本实现了之前技术无法达到的规模。这些技术使得人们可以处理的数据量大大增加，并且不再需要用传统的数据库表格来整齐地排列。它们奠定了当前大数据技术的基础，可以认为是大数据技术的源头。

大数据应用首先出现在互联网、电信、金融行业。一方面源于这些行业在数据积累方面具有得天独厚的先天优势，另一方面更在于其业务运营自身就带有“数据驱动”色彩。从“搜索引擎”起步，到智能推荐、定向广告服务，直到今天，几乎所有的互联网业务都与大数据技术密不可分；电信、金融行业也最早利用大数据技术诊断业务，改善服务水平，以获得更多的客户和更高的业务增长。而随着传统行业的数据化，大数据将得到更为广泛的应用。

人类社会已经进入“大数据”时代，这部分归功于物联网、无处不在的

1
<http://webmarketingtoday.com/articles/ad-anorexia/>

2
<http://www.google.com/about/company/history/>,
2008年7月谷歌搜索引擎能检索的网页数突破1万亿个

终端和传感器，部分归功于依靠大数据精确分析与预测的商业模式以及使这一切成为可能的强大算法设计^[2]。

1.2 大数据的含义

大数据的概念最早由全球知名咨询公司麦肯锡提出，麦肯锡认为，数据，已经渗透到当今每一个行业和业务职能领域，成为重要的生产因素。人们对于海量数据的挖掘和运用，预示着新一波生产率增长和消费者盈余浪潮的到来。维基百科将大数据定义为一个复杂而庞大的数据集。香山科学会议则认为大数据是来源多样、类型多样、大而复杂、具有潜在价值，但难以在期望时间内处理和分析的数据集。国际数据公司（IDC）归纳了大数据的4个特征（4V特征），即海量的数据规模（Volume）、快速的数据流转和动态的数据体系（Velocity）、多样的数据类型（Variety）和巨大的数据价值（Value）³。

认识大数据，要把握“资源、技术、应用”3个层次。大数据是具有体量大、结构多样、时效强等特征的数据；处理大数据需采用新型计算架构和智能算法等新技术；大数据的应用强调以新的理念辅助决策、发现新的知识，更强调在线闭环的业务流程优化。因此说，大数据不仅“大”，而且“新”，是新资源、新工具和新应用的综合体⁴。

1.3 大数据应用发展

“大数据”的发展与IT产业其他领域的发展相辅相成，近年来互联网、移动互联网、物联网、云计算和高性能计算等方面的高速发展从内涵上不断推动大数据的技术演进，从外延上不断延展大数据的应用范围^[3]。大数据已经撼动了实际的方方面面，从商业科技到医疗、政府、教育、经济、人文以及社会的各个领域。

整体而言，大数据应用在各行各业的发展呈现“阶梯式”格局：互联网行业是大数据应用的领跑者，金融、零售、电信、公共管理、医疗卫生等领域积极尝试大数据。大数据让各行各业收集以往无法想象的海量数据，并帮助人们深入挖掘其中的价值，对国家、企业和个人而言都具有积极意义^[4]。

（1）互联网行业是大数据应用的领跑者

互联网是大数据应用的发源地，大型互联网企业是当前大数据应用的

3
The 2011 Digital Universe
Study: Extracting Value
from Chaos, IDC 2011

4
大数据白皮书（2014）.
工业和信息化部电信
研究院

领跑者。搜索引擎是最早的互联网大数据应用，谷歌在2000年左右提出了MapReduce、BigTable等技术，从此开启了大数据技术的新篇章。经过十多年的发展，互联网上形成了多种相对成熟的大数据应用模式，按照用途可分为以下3类。

第一，商业大数据应用，即主要以盈利为目的的大数据应用。目前常见的应用有：一是基于用户个人信息、行为、位置、微博等数据而进行的个性化推荐、交叉推荐、品牌监测等营销类大数据应用。由于其商业模式清晰可见，市场需求广泛旺盛，因此这是目前互联网上最热门、最普遍的应用，被互联网广告、电子商务、微博、视频等公司普遍采用。二是基于用户、商铺的交易数据而进行的经营分析报告、反欺诈、反虚假交易、促销和团购选品、产业集聚判断等交易辅助类大数据应用，这些应用目前已经逐渐成为电子商务企业的必备工具。三是利用网站动态数据对网络状态进行实时监控预警、网站分析优化和网络信息安全保护（如用户信息保护、防网络攻击、防钓鱼、防垃圾邮件等）的网络安全大数据应用。

第二，公共服务类大数据应用，即不以盈利为目的、侧重于为社会公众提供服务的大数据应用。典型案例如谷歌开发的流感、登革热等流行病预测应用，能够比官方机构提前一周发现疫情爆发状况。国内也有搜索引擎公司提供诸如春运客流分析、失踪儿童搜寻的公益大数据服务。

第三，技术研发类大数据应用，即利用大数据技术促进前沿技术研发、持续改进产品性能的应用。互联网应用在新版本的研发中，常常进行A/B测试，就是大数据在产品开发中的典型应用。A/B测试中，服务商同时收集新老版本下的用户行为数据（如点击行为、访问时长、鼠标停留等），并进行分析比对，用以指导产品后续的改进方向。另外，利用各种语言版本的网页数据不断提高翻译质量的机器翻译、利用更多语音指令不断提升质量的语音识别技术以及无人汽车等前沿技术的研发，也广泛应用了大数据技术⁴。

（2）电信、金融行业积极跟随

根据全球电信行业大数据分析市场（2014—2018年）报告，电信领域对于大数据分析技术工具的应用将以28.28%的年度复合增长率快速发展。运营商积极利用其掌握的网络运营数据、用户行为数据，开辟新的数据变现路径。

在金融行业，大数据分析帮助银行、保险机构更好地细分用户群体，提供个性化服务，降低风险控制成本，更催生了新兴的互联网金融服务。

(3) 传统行业加快大数据应用

根据中国信息通信研究院 2016 年发布的《大数据白皮书》调查显示：相比于互联网、电信、金融行业，我国传统行业的大数据应用发展相对滞后。例如：批发零售业超过 80% 的企业仍没有形成大数据应用计划，远低于大数据应用的整体平均水平。然而可预见的是，随着物联网的加速部署，数据采集渠道不断拓宽，政府数据开放运动的深入推进，大数据分析技术在传统行业和领域必然呈现快速渗透的趋势。在交通、医疗、教育、能源、工业制造甚至农业领域得到普遍应用。美国的传统农业服务企业孟山都，通过运用大数据技术，成功转型为一家农业大数据科技公司就是一个典型的例子。该公司在其本身所积累的育种、农药、化肥等农业数据基础上，通过收购整合与农业服务相关的数据，如气象数据、土壤数据，以及通过服务方式免费获得农民种植农作物的生长数据，建立了全新的农业大数据商业模式。孟山都的农业大数据服务为农民提供种植信息服务，农民可以通过软件实时监测气候及土壤肥力变化，实现差异化播种、灌溉等种植活动。与此同时，孟山都通过应用大数据分析，培育更加具有竞争力的农作物，更为重要的是，孟山都通过分析农作物生长周期数据，能够更加主动把握大宗农业商品价格变化，为农民提供更为精准的保险服务。

大数据发展的拐点已经到来，目前正在逐渐成为经济活动的主要承载者。数据被称为资产、原油、原材料、货币，无论哪种形容的方法都不过分。IDC 预测：2020 年，70 亿人的数据化生存以及 500 亿个互联设备的感知、互联和智能，将产生 35 ZB 的数据。1 ZB 相当于 1 000 EB，目前谷歌公司的数据量级为数十 EB，这就意味着，一年将产生相当于 1 000 个谷歌公司的数据量^[3]。

据统计，2014 年，全球大数据市场规模达到约 285 亿美元，实现 53.23% 的增长。2015 年，全球大数据总体加速发展趋势不变，全球大数据市场规模将达到 421 亿美元，预计 2020 年全球大数据市场规模将达到 1 263.21 亿美元，同比增长 17.51%。另一方面，网络数据成为重塑国家竞争优势的新动力。大数据被誉为“21 世纪的石油和金矿”，已经成为与物质、能源同等重要的国家基础性战略资源。

世界各国高度重视数据资源战略价值，相继出台国家战略，落实配套措施，系统推动大数据发展。美国带头三步走，2009 年 Data.gov 网站上线，开放政府数据，做大数据资源体量；2012 年，美国白宫发布大数据研发计划，

资助研发、带头应用；2015年，美国白宫发布白皮书《抓住机遇，守护价值》，总结大数据中隐私保护政策。各国紧随美国战略布局。2013年，英国发布《英国数据能力发展战略规划》，日本发布《创建最尖端IT国家宣言》，澳大利亚发布《公共服务大数据战略》^[4]。我国也在2015年出台了《关于运用大数据加强对市场主体服务和监管的若干意见》（国办发〔2015〕51号）和《国务院关于印发促进大数据发展行动纲要的通知》（国发〔2015〕50号）等一系列促进大数据应用与发展的国家政策文件。

1.4 大数据带来隐私风险

在全球信息经济中，个人数据已经成为在线活动的第一驱动力。伴随而来的隐私与个人数据保护、信息安全问题日益受到关切。数据保护与数字经济中的货物和服务贸易直接相关。不充分的隐私保护会对产品和服务带来直接的消极影响，损害消费者信任度，当然过于严格的数据保护也可能会对贸易带来负面影响⁵。

随着大数据挖掘分析越来越精准、应用领域不断扩展，个人隐私保护和数据安全变得非常紧迫。在大数据交易中的许多标的都是基于以个人为粒度的数据，即使这些数据经过了一定的清洗，但也很难保证个人的隐私不被泄露。

并不是所有的大数据应用都涉及个人信息。例如工业互联网中采集的工厂生产数据、气象数据、国家地理测绘数据等都不包括个人信息。大数据的数据类型，从不同的视角出发，可以有不同的划分，例如可以包括：Web与社交媒体数据、机器对机器的数据、大体量交易数据、生物计量学数据、人工生成的数据。从行业的视角出发，可以看到医疗行业的数据、公共事业部门的数据，工业、农业领域的数据^[5]。但不容否认的是，大数据中采集的相当一部分数据包含有个人信息，特别是目前阶段的大数据应用更是主要集中于对个人数据的分析加工。

大量电子、计算机设备的部署，让人们的行为数据得以记录，成为大数据分析的原料。人类正处于日趋精密的被监视状态之中。事实上，个人的一举一动都能在某个数据库中找到线索。

电子邮件都保存在邮件服务商的日志文件中；通话记录都被加上时间标记备份在电话公司的大容量硬盘之上，人们何时何地买了什么东西，人们的喜好、品位以及支付能力都被信用卡提供商编目归档；用户所有的个人网页、空间、Facebook文件，还有博客信息都被保存在多个服务器上；用户的即时

5 United Nations Conference on Trade and Development (UNCTAD), Data protection regulations and international data flows: implications for trade and development

行踪完全被移动服务商掌握。人们的容貌和穿着打扮都被安装在各大商场和街角的摄像头捕捉并记录。虽然通常选择不去多想，但事实上，人们的生活完全能被这些雨后春笋般出现的数据库所记录的信息串联起来。

这些记录引爆了个人隐私危机，这一问题的严重性再怎么夸大也不为过。然而，它同时也创造了一个历史机遇，第一次毫无偏见地提供了成千上万人，而不是一部分或者少数人的详细行为记录。企业利用这些数据来优化生产、服务，政府利用数据来追捕犯罪分子，无数个新兴企业想利用数据追踪个人的一举一动，从中寻觅商机，期望变成下一个谷歌^[6]。在人类历史上，数据第一次被大规模地采集、分析，并用于对未来的预测，个人的行为数据被企业、政府仔细研究分析和研究，为了开发一个更受欢迎的产品，抑或发现潜在的恐怖分子，无数的计算机和研究专家正在集中精力预测人类行为。数据挖掘把数据分析的范围从“已知”扩展到“未知”，从“过去”推向了“将来”^[7]。

赤裸裸的未来会给人们带来更安全的城市、更聪明的学生、更优秀的电影、更健康的身体和更广阔的事业，而同时，也增添了新的沮丧、不便和不公平，人们会感到被侵犯。这就是大数据应用带来的两面性。换言之，在大数据时代，数据共享的价值或福利是集体体验，而风险却是由个人承担的。

以大型零售百货商塔吉特（Target）与怀孕指数为例进行说明。在未经得用户的同意下，经过算法得出了个人信息，甚至是敏感的个人敏感信息。这在 Target 的案例中体现得十分明显。Target 是美国最大的连锁超市之一，该公司使用数据挖掘极其有效地提高营销精准率，能做到在事情显现之前就预测到它的发生。

该超市基于数据挖掘系统分析结果给一位美国高中女生寄去婴儿用品优惠券，其父亲发现后投诉 Target 误导未成年人，但却在之后了解到他女儿已经怀孕的事实。原来 Target 超市的数据库系统给每个顾客分配一个 Target Guest ID，在该条目下详细记录顾客的信用卡信息、网上注册信息，在 Target 官网浏览的每一个页面和停留时间长短、每次的购买行为等信息。数据挖掘团队专门分析这些收集到的历史信息，预测顾客将来的购物行为和需求甚至生活方式，然后发邮件给顾客。

例如判断一个女性怀孕，线索是该顾客已经发生了的消费行为，她可能购买了维生素、大量的专用乳液、无水洗手液等典型的孕妇会购买的一些商品，在孕妇、婴儿用品页面停留较长时间等。由于美国人非常注重隐私，为了避免上述早于父亲发现女儿怀孕这种尴尬事件发生，Target 针对性地改变营销策略，把母婴系列产品的优惠券和信息混合在其他产品的信息里发给顾客，掩人耳目，