HZ BOOKS
华章IT

大数据管理丛书

孟洁　薛颖　朱玲凤　著

数据合规

入门、实战与进阶

企业数据合规实用手册　　数据合规律师成长指引

机械工业出版社
China Machine Press

大数据管理丛书

数据合规：入门、实战与进阶

孟洁　薛颖　朱玲凤　著

ISBN：978-7-111-70536-9

本书纸版由机械工业出版社于2022年出版，电子版由机械工业出版社华章分社出品，授权北京世纪卓越信息技术有限公司在全球范围内制作与发行。

客服热线：+ 86-10-68995265

客服信箱：service@bbbvip.com

官方网址：www.hzmedia.com.cn

新浪微博 @华章数媒

微信公众号 华章电子书（微信号：hzebook）

目录

作者简介

孟洁

现任北京市环球律师事务所合伙人，主要执业领域为网络安全、个人信息与隐私保护。曾在多家知名企业担任法务负责人和数据保护官，任IAPP中国区知识社区主席，被钱伯斯、The Legal 500、LEGALBAND等知名法律评级机构评为“TMT领域领军人物”“数据保护领域领军人物”“Fintech领域头部律师”等，被北京市律协评为全国千名涉外专家律师。

薛颖

长期在互联网集团担任数据合规与知识产权总监。在外企、世界五百强公司等从事过多年数据隐私合规工作，拥有丰富的互联网场景一线经验。持有CIPP/E、CIPP/U认证，当选ALB中国知识产权法务15强并带领团队获得过《商法》年度“数据合规”优秀团队等奖项。

朱玲凤

现任知名互联网公司隐私及数据合规专家，曾任小米安全与隐私委员会隐私副主席。多年从事数据隐私合规研究和实务工作，深入参与国内信息安全相关标准拟定和重要法律研讨等，在全球隐私法律研究、隐私保护设计、隐私安全技术应用与管理以及App、物联网、人工智能等领域有丰富的实践经验。

自序

掐指算来，这些年我们作为企业法务和律师，其中很长时间都是从事一线数据合规工作，加起来也有将近20年的“数据合规”工作经历了。那时我们都还年轻（当然，现在热爱学习的心也依然），对这个领域无知且无畏，那时没有《网络安全法》，也没有GDPR（暴露年龄了），那时我们看到数据处理条款和“同意”要求都还很懵圈，那时我们作为中国律师在这个领域没有发言权，需要跟随域外法律从头学起……时光荏苒，一晃已经过去了好多年。

后来我们不知不觉在这个领域一路走下来，不断成长，也不断收获，平时对一线数据合规治理工作有了些心得体会，就赶紧写下来。细壤不拒，细流不择，慢慢有了最初10万字的积累，再后来就有了体系化、做“成书”，以总结传承的想法。然而，每每汇总整理时，却又都重重受阻，不是新法更新太快，就是日常工作太忙。其实，归根结底还是内心忐忑，不知成书是真的能够帮助大家，还是贻笑大方。

赶上我国《个人信息保护法》的制定、颁布和生效，也算是从事数据合规工作遇到了一个里程碑。我们决定以此为动力，督促自己一定把这本书最终写完。

本书将以一位数据合规法务或律师“白晓萌萌”的成长之路为脉络，分别从入门篇、进阶篇、高阶篇逐步介绍数据合规领域专业人士一路成长过程中会遇到的各类业务场景和风险点，探讨梳理各场景下的数据合规治理解决方案，为希望了解、从事或喜欢数据合规这个领域的法务或律师们拨开云雾，提供数据合规治理的门径与指引，并展望这一专业领域的职业前景和蓝图规划。

写这样一本书并不是任务，也没有指标，更无关奖酬，就是为了提醒、督促、鉴证自己在数据合规这个枯燥而又生动的领域不懈努力、不忘初心。其实说起来，在这个领域坚持下来的初心真就以下两个，虽看起来有些太过“诗与远方”，却不惮于读者诸君窃笑无知狂妄而分享于此。

一是从小处着手，希望在一线业务场景中实现“Law is Code”。多年前初见劳伦斯·莱斯格教授的金句“Code is Law”，并无甚感觉，然而在一线互联网或物联网场景从事隐私保护设计工作多年下来，对此经典判断颇感认同，并有了进一步的化用和体会，还希望“Law is Code”，乃至“Code is Code”——在应然状态下，良好的“隐私保护设计”方式确保软件代码与法律规范要求（法律侧代码，另一种“Code”）一致，让技术和法律两种“Code”协同实现业务功能，落实法律对个人信息与数据保护的要求。例

如，在用户做出“同意”之前，不能触发SDK来收集用户个人信息，避免不必要地高频读取用户终端地理位置，可以便利地关闭App的访问权限……这些细致具体的数据合规工作落地到一线场景，就是需要让法律规范要求限制技术代码，而不是让技术代码自行其是地获取数据，做出任意处理；同时，也要借助技术手段来高效实现法律对个人信息和重要数据等法益的保护。

这些工作初始可以自己做，随后带动团队做，再后来多方一起努力，影响一条业务线、一家企业，甚至一个行业参与其中。写书不失为一种有效的方式——影响更多人关注和投入这个领域。

二是从大处着眼，希望能为推动数据治理规则更加完善献上一份微薄之力。数据合规的治理规则是数字化时代全社会治理规则体系不可缺失，甚至越来越重要的组成部分。时下热议的数字经济、人工智能、元宇宙等的发展背后都离不开数据伦理和数据处理规则。良善治理和共同富裕的美好社会使普通人受益于数据使用带来的安全、便利和更多福祉，同时，亦无须无限制地让渡个人信息保护和隐私，当然也会避免由资本裹挟技术推动数据使用的无限扩张。

只有在法律的治理与规则的约束下，个人信息、重要数据、大数据和人工智能等与数据相关的要素和资源才会被确保用于做“好”的事情。在AI还不能自主生成法律规则之前，我们还可以抓紧时间探索和制定愈加成熟的、推动技术向善发展的规则体系。如果说，在以前这是重要而不迫切的事情，那么随着个人信息收集使用、大数据算法和人工智能越来越广泛运用，这件事已经变得重要且迫切了。正如苹果公司CEO库克在2021年1月的一次演讲中提到的：“如果我们接受生活中的一切都可以被汇总和出售，并且认为这是正常的、不可避免的，那么我们失去的不仅仅是数据，而是失去了做人的自由。”

当然，为个人信息保护和数据使用设立规则并非为了阻碍数据的利用，而是为了促进数据被合法有效地利用，真正让个人、企业和社会都能得到保护，实现各方利益共赢，达到并保持“有效保护与合法使用的可持续状态”。

近代以来的法治精神始终贯穿着“尊重人之为人”的思想，如果法律人能够做一点点事情，推动完善数据治理的规则，以避免把人异化为可以被随意处理的数据字段，使得大数据时代下的人仍然可以保有隐私与尊严，维持人格独立，享有思想自由，这也算是法律人在大数据和人工智能时代回应时代命题的价值和使命吧。

在极其忙碌的日常工作之外，还要坚持不懈地自虐码字，又不揣粗陋成书出版，只愿为我们这个时代的数据保护和治理水平之提升尽绵薄之力。无

论是解决一个个具体的代码场景，还是有利于整个社会的数据治理规则完善，我们作为有幸处在这个变革时代的个体，感受着“数据”为我们的生活带来的变化，总想着要为此做些事情。

星辰大海虽远，然心向往之并孜孜以求，正是“怕什么真理无穷，进一寸有进一寸的欢喜”。

是以为序。

作者

2022年1月 北京冬日暖阳

开篇　小白入职“数据合规”法务岗位，一头雾水怎么办

- 第一章　“数据合规”都管哪些事儿

- 第二章　数据合规之避坑预警

白晓萌萌就职于一家互联网公司法务部，公司的主营业务是面向C端用户提供在线内容服务，涉及在线音视频、在线教育、社交等业务线。在被总法律顾问指定来做“数据合规”专职律师的时候，她整个人还没有完全反应过来。她知道这个领域非常前沿，知识迭代快速，尚处于摸索和创新阶段，因此感到不知如何下手，从哪里切入。当然，她还不知道的是，总法律顾问指定她来担任这个岗位，不是因为她懂这个领域——事实上，大部分人都不懂。领导相信她这个“小白”能够成长为专家，乃是因为白晓萌萌极强的学习能力和经验迁移能力，以及对互联网产品开发设计的热爱，还有对“数据合规”这个领域的好奇。

本篇就是帮助白晓萌萌这位“小白”律师从零开始了解自己的岗位职责。开展数据合规一线工作，首先要弄明白以下问题：

- 要管“哪些数据”——哪些“数据”受到法律和监管要求，需要合规？

- 要管数据的哪些事儿——哪些与“数据”有关的业务运营活动需要“合规”？

- 如何做数据合规——数据合规岗位通常需要做哪些事？会与哪些其他岗位产生工作交集和相互配合？

- 数据合规有哪些底线——这个领域最不能触碰的“大坑”有哪些？如何练就避坑大法？

理解数据合规领域的基本范畴、客体、岗位职责、合规红线——这些都是入门基础知识，只是这个“入门”却不那么容易掌握，虽是“起点”，却也是数据合规领域的“难点”之一。一旦基础掌握不好，以后小白如遇到各种涉及数据处理活动的业务场景，动辄阵脚大乱、动作“变形”或“跑偏”，出具的法律意见和解决方案如无本之木、无水之源，或无法落地，或无从满足法律要求，或根本难以控制风险。

千里之行始于足下，白晓萌萌从零开始，正式开始她的数据合规律师升级打怪之路啦！

第一章　“数据合规”都管哪些事儿

【场景】白晓萌萌最近遇到了入职以来的第一次大挑战：不知道领导安排的新岗位是做啥的。

公司近期频频因个人信息保护相关事宜被通报整改，受到行政处罚，引发民事纠纷等，业务经常因行政处罚、整改要求等需要紧急调整产品交互、后台设置等，否则将面临应用商店下架、整个业务停滞的风险，影响业务整体的规划实施。然而，公司有些部门有时候甚至连合规要求是什么都不了解，整改无从下手，叫苦不迭。公司领导提出希望由专人来管理数据合规，梳理出明确的合规要求，与整个产品开发流程相结合，提前防控风险。总法律顾问指定白晓萌萌主要负责个人信息保护事宜。

白晓萌萌可谓数据合规领域的“小白”律师，因为第一个问题就难住了她：“数据合规”都管哪些事儿呢？

第一节　这些数据很重要：用户数据、个人信息、隐私

白晓萌萌首先需要界定数据合规管的到底是哪类数据。她对相关业务部门进行了访谈，了解到企业内部有各种各样的数据，包括：企业经营数据，如财务报表、现金流水、产品日激活人数和活跃人数；企业决策所需数据，如行业统计报告；产品收集的各类数据，包括用户的注册信息、行为信息等；企业在收集的各类数据基础上加工开发的数据，如用户画像、推荐算法模型、产品优化方向等。访谈后，白晓萌萌得出结论：数据合规管的是与用户相关的数据，具体边界并不清晰，而且用语都不同，有的人称之为用户数据，有的人称之为个人信息，大多数人称之为隐私。

白晓萌萌检索了相关的法律规定、国家标准后了解到，海外某些国家会使用“personal data”这个名词，翻译成中文即“用户数据”，而我国更多使用的是“个人信息”，这二者的所指是一致的，但是在各国的法律规定下，其边界可能不一致，需要根据各国法律的规定来界定。因本书基于国内法，故统一为“个人信息”。

1.什么是个人信息

个人信息是以电子或者其他方式记录的与已识别或者可识别的自然人有关的各种信息，不包括匿名化处理后的信息[1]。如姓名、出生日期、身份证件号码、个人生物识别信息、住址、联系方式、通信记录和内容、账号密码、财产信息、征信信息、行踪轨迹、住宿信息、健康生理信息、交易信息等[2]。

根据上述定义，个人信息能够识别出“特定”个人即可，无论是识别出该特定自然人的身份（你是谁），还是只识别出一个自然人的行为活动（做了什么）、终端设备（用的什么）等。因此，只要借由信息本身的特殊性可以将某个特定自然人与其他人区分开来，完成识别（即使不知道该自然人的姓名和社会身份），该信息就构成了“个人信息”。

通过个人信息识别出自然人既可以是直接识别，如通过身份证号、护照号码等身份硬标识符自身直接识别出特定自然人；也可以是与其他信息结合来识别出特定自然人，如通过结合移动设备识别符与网络浏览历史，区分出特定自然人，进而向其提供个性化的新闻列表等。

从现实生活的角度，一般通过身份证号、护照号码或者人脸识别、指纹识别等生物识别方式来识别人，其核心在于与自然人的唯一且密切的关联性。从通信产业的角度，识别人的方式是手机号码。但进入虚拟的网络世界后，标示一个用户的最常见方式则是虚拟账号，如微信号、淘宝账号等。互联网服务商通过账号来识别用户，再通过注册账号的手机号、邮箱等联系方式向用户发送广告等。此时，识别用户不一定要知道用户姓甚名谁，只需要唯一关联到某个人即可，因此，相比现实世界的识别符来说，这种识别方式与自然人的关系没有那么密切。

随着移动互联网的发展，相比账号和手机号码来说，移动设备识别符这一对于普通用户而言更无法感知和理解的信息在识别用户、形成用户画像、提供个性化服务等方面起到了重要的作用。用户在移动互联网时代经常会遇到此类场景：在淘宝上搜索健身器材，到今日头条上就能看到同类的广告，甚至在微信或电话中谈到旅游，在抖音中就看到了同类的广告。是App有千里眼和顺风耳吗？当然不是，上述场景实现的关键是依靠移动设备识别符以及用户画像技术。

移动设备识别符是指手机或者其他移动设备上唯一标识该设备的识别符，类似于该移动设备的身份证号码，如MAC地址、产品序列号（SN）、Android系统上的IMEI、iOS系统上的IDFA。移动设备识别符在识别用户方面比账号的作用更大，账号一般来说仅是一个公司或者一个产品内识别同一个用户的媒介，而移动设备识别符是同一个移动设备上所有服务商都认识同一个用户的媒介，广告主可以通过移动设备识别符定位出不同公司账号对应的同一个用户，于是就发生了各App之间仿佛认识同一个用户的情况。

因此移动设备识别符也是一种非常重要的个人信息，即使它本身并不能识别出该特定自然人的身份，但可识别出特定自然人使用的终端设备，从而识别出一个自然人的行为轨迹，进而完成精准的画像。

为了便于大家理解，《信息安全技术　个人信息安全规范》（GB/T 35273—2020）中对个人信息进行了举例，如图1-1所示。

敏感个人信息是一旦泄露或者非法使用，容易导致自然人的人格尊严受到侵害或者人身、财产安全受到危害的个人信息，包括生物识别、宗教信仰、特定身份、医疗健康、金融账户、行踪轨迹等信息，以及不满十四周岁未成年人的个人信息[3]。值得注意的是，不同国家对于敏感个人信息的范围认定存在差异，而且可能使用类似“特殊类型的个人信息”的概念。

敏感个人信息的滥用或泄露将造成更为严重的风险，因此在合规要求上会更为严格，包括在具有特定的目的和充分的必要性，且采取严格保护措施的情形下[4]，并在额外告知处理敏感个人信息的必要性和对个人权益的影响后经过个人的单独同意等[5]。

个人基本资料	个人姓名、生日、性别、民族、国籍、家庭关系、住址、个人电话号码址等
个人身份信息	身份证、军官证、护照、驾驶证、工作证、出入证、社保卡、居住证等
个人生物识别信息	个人基因、指纹、声纹、掌纹、耳廓、虹膜、面部识别特征等
网络身份标识信息	个人信息主体账号、IP 地址、个人数字证书等
个人健康生理信息	个人因生病医治等产生的相关记录，如病症、住院志、医嘱单、检验报告记录、护理记录、用药记录、药物食物过敏信息、生育信息、以往病史、诊病史、现病史、传染病史等，以及与个人身体健康状况相关的信息，如体重量等
个人教育工作信息	个人职业、职位、工作单位、学历、学位、教育经历、工作经历、培训记
个人财产信息	银行账户、鉴别信息（口令）、存款信息（包括资金数量、支付收款记录等信贷记录、征信信息、交易和消费记录、流水记录等，以及虚拟货币、虚拟兑换码等虚拟财产信息
个人通信信息	通信记录和内容、短信、彩信、电子邮件，以及描述个人通信的数据（据）等
联系人信息	通讯录、好友列表、群列表、电子邮件地址列表等
个人上网记录	指通过日志储存的个人信息主体操作记录，包括网站浏览记录、软件使用录、收藏列表等
个人常用设备信息	指包括硬件序列号、设备 MAC 地址、软件列表、唯一设备识别码（如 ID/IDFA/OpenUDID/GUID/SIM 卡的 IMSI 信息等）等在内的描述个人常用的信息
个人位置信息	包括行踪轨迹、精准定位信息、住宿信息、经纬度等
其他信息	婚史、宗教信仰、性取向、未公开的违法犯罪记录等

图1-1　《信息安全技术　个人信息安全规范》附录A中对个人信息的举例

2. 个人信息与隐私辨析

隐私是自然人的私人生活安宁和不愿为他人知晓的私密空间、私密活动、私密信息[6]，权利主体行权的核心目的在于“不愿为人知晓”和“生活安宁”，即隐私的核心是“私密性”，即信息主体不想让外界知悉这些空间、获得的信息，即使有些信息并不属于他的个人信息，甚至只是一种物理空间（私人活动所覆盖的范围）。由于主体行权的目的多在于防御第三人获知，因此是一种相对偏主观感觉的认知。个人信息如上所述，核心是“可识别性”，即该信息可直接或结合识别出特定自然人，主体行权目的多在于想要积极管理已授权第三人处理的个人信息。但是否属于其个人信息，则是相对偏客观事实的认定。因此，针对自然人的隐私或个人信息而言，“私密性”和“可识别性”是两种重要但不同的划分维度，存在一定程度上的重合（即主观上不愿让外界知悉的、客观上可识别其个人的信息）。表1-1所示为个人信息与隐私的辨析。

表1-1　个人信息与隐私辨析表

情形	信息的二元属性	
1. 既是隐私，又是个人信息	既有私密性，也有可识别性	基因数据
2. 不是隐私，仅是个人信息	无私密性，但有可识别性	面部特征
3. 仅是隐私，非个人信息	有私密性，无可识别性	将游戏人
4. 既非隐私，也非个人信息	无私密性，无可识别性	匿名化的

值得注意的是，表1-1的前三种情形存在此消彼长、相互角力的有趣情况。特别是，随着大数据时代进行信息收集、追踪、匹配的能力越来越强，个人信息的边界在不断扩展，而可被归入隐私范畴的信息却越来越少，即情形2扩张为主流样态，情形1和3的空间却越来越小[7]。这才导致人们常常有一种错觉，认为个人没有隐私，甚至将个人信息与隐私混同。

如前所说，作为隐私权和个人信息重叠保护对象的“私密信息”是《中华人民共和国民法典》（以下简称《民法典》）下一个重要的新生术语，也是最难界定的人格权客体之一。只有一项信息同时具备“能够识别自然人”和“主体不愿为他人知晓”两个要件，才是构成私密信息的个人信息。

但在个人信息处理和司法实践中，对某一信息属性的认定难点往往不在于“可识别性”，而在于难以判断“私密性”，即构成个人信息的同时是否还会进一步构成私密信息。然而，是否构成私密信息不仅是一项主观认定，还需要同时兼顾考虑三重视角：信息主体个人的合理隐私期待，即作为信息主体个人的内心感觉、想法、诉求、期待、愿景；普通人的一般合理认知，即作为一个普通人通常会做出的预期判断与根据正常人的社会经历和认知水平会做出的内心反应；实际的信息收集和处理场景，即在特定情况下并结合当时的具体情境会做出的心理判断与暗示。

如果以上任一因素发生变化，对于个人信息的私密性判断都可能出现不同结果。例如，极端来讲，即使性取向系敏感个人信息，公众亦一般认知为隐私，但是否绝对属于隐私，也要看主体自身的性格、认知、理解力与抗压性，主体所在环境的包容性与开放性，是否有主动意愿公开，是否为实现特定诉求，是否因此会遭受歧视性待遇等。例如，在一些较为传统的社区环境，个人一般不愿意公开同性取向，而在开放程度大的城市，愿意公开同性伴侣关系的人越来越多，甚至一些公众人物公开表明自己的性取向，就更难以构成其隐私了。当发生隐私权与个人信息认定的争议时，法律应当同时考虑以上三点，不可偏废一端，才能保证其公平公正。

在北京互联网法院的“微信读书”案中，判决从用户合理隐私期待的维度将个人信息兼具的隐私属性划分为三个层次：一是符合社会一般合理认知下共识的私密信息；二是不具备私密性的一般信息；三是兼具防御性期待及积极利用期待的个人信息。是否侵权需要结合信息内容、处理场景、处理方式等进行符合社会一般合理认知的判断[8]。另一民事判决从场景化角度来分析，指出社交应用的好友关系在一定范围内已公开，并非不愿为他人知晓的私密信息。当然，这些诉争标的兼具个人信息与隐私权争议的案件还比较零星，尚未形成普遍共识。虽然对于社交软件中的虚拟社交关系、电话通讯录等构成个人信息的分歧不大，但这些信息是否同时仍然具有“私密性”的属性，则在企业和用户之间产生很大分歧。不同社交软件对社交关系的公开模式多有差异，因此用户的合理隐私期待也应该场景化识别，即不应脱离具体软件的适用场景概括出过于抽象的论断。

3. 个人信息与匿名化、去标识化的辨析

此外需要辨析的概念是：部分信息确定是个人信息，但是可能进行一定的加工处理，对其识别性产生了影响，那么处理后的个人信息是否仍然属于个人信息呢？还有一个需要辨析的概念是匿名化，《中华人民共和国个人信息保护法》（以下简称《个人信息保护法》）规定，“匿名化是指个人信息经过处理无法识别特定自然人且不能复原的过程”[9]。如果认定为符合匿名化的要求，则匿名化处理后的信息不再适用个人信息处理的法律要求，因为其已经不再具备可识别性。但是，需要注意的是，对个人信息匿名化的处理要求是比较高的，需要无法从处理后的结果复原、识别特定自然人，而且应持续采取相应的机制来防范随着技术发展或者数据融合而从中重新识别出特定自然人的风险。

随着快递物流、外卖等场景下手机号码被泄露情况的频繁发生，日常生活中出现了以“微笑面单”方式（将用户手机号码中间四位数隐藏为笑脸符号或星号），及外卖平台或者网约车平台提供“隐私号”等对手机号进行“脱敏”的处理方式。但是这是否意味着隐藏中间四位号码的手机号码或者替换为唯一值等处理后的信息就不属于个人信息了呢？事实上，在绝大部分情况下，这样简单脱敏处理的手机号并没有匿名化，仍然是个人信息。

手机号码经过简单隐藏四位数或者替换为唯一值的隐私号等处理后的信息通常并不能构成上述匿名化要求。对于进行上述处理的企业而言，仍会保留处理前后信息的映射表，以持续识别用户以及提供相应服务。因此，此类简单脱敏处理的手机号不属于无法识别特定自然人且不能复原；对于接收该等手机号的主体而言，通过将不同渠道获取的手机号或其他信息进行“撞库”匹配，也存在再次还原出原始手机号的可能性。

所以，简单隐藏手机号码的几位数或替换为唯一值的隐私号等处理方式，其实仅仅是去标识化而已——去标识化是“个人信息经过处理，使其在不借助额外信息的情况下无法识别特定自然人的过程”[10]。借助隐藏的位数或者数据映射表，还是能够从隐藏四位数的手机号或隐私号识别特定自然人的。这样去标识化的信息并不意味着不再属于个人信息，无须按照个人信息保护相关规则加以处理，而仅是个人信息处理的安全保障措施之一，用以避免在类似快递单、外卖订单等个人信息展示环节泄露个人信息，以及违法人员在接收到个人信息后可以快速、便利地定位到特定的自然人而进行诈骗等。

[1] 《个人信息保护法》第4条第1款。
[2] 《信息安全技术　个人信息安全规范》（GB/T 35273—2020）第3.1条以及附录A。
[3] 《个人信息保护法》第28条第1款。
[4] 《个人信息保护法》第28条第2款。
[5] 《个人信息保护法》第29条和第30条。
[6] 《民法典》第1032条。
[7] 在“庞先生诉某航空公司”一案中，二审法院明确指出：“在当今的大数据时代，信息的收集和匹配成本越来越低，原来单个的、孤立的、可以公示的个人信息一旦被收集、提取和综合，就完全可以与特定的个人相匹配，从而形成某一特定个人的详细而准确的整体信息。这些整体信息一旦被泄露扩散，任何人都将没有自己的私人空间，个人的隐私将遭到巨大威胁，任何他人未经权利人的允许，都不得扩散和不当利用能够指向特定个人的整体信息。”
[8] 北京互联网法院（2019）京0491民初16142号判决书。
[9] 《个人信息保护法》第4条、第73条第（四）项。
[10] 《个人信息保护法》第73条第（三）项。

第二节　要管理的数据处理活动太多了：覆盖数据全生命周期

白晓萌萌在厘清了数据保护覆盖的对象为个人信息后，需要进一步了解数据合规工作的范围。从信息技术的本质而言，个人信息是一个或几个字段，即数据。如“01010202,F,click,2021-04-21 9:26:00”，该行数据根据自定义的数据结构，含义为“ID是01010202，性别为女，在2021年4月21日9点26分发生了一次点击行为”。数据有自己的生命周期，如图1-2所示[1]，从逻辑上可以简单分为数据收集、使用、存储、披露至销毁。使用“数据全生命周期”的框架，一方面符合数据的基本规律，另一方面可以帮助数据合规人员全面梳理企业处理个人信息的活动，进而分阶段评估和处置相应的个人信息保护风险。

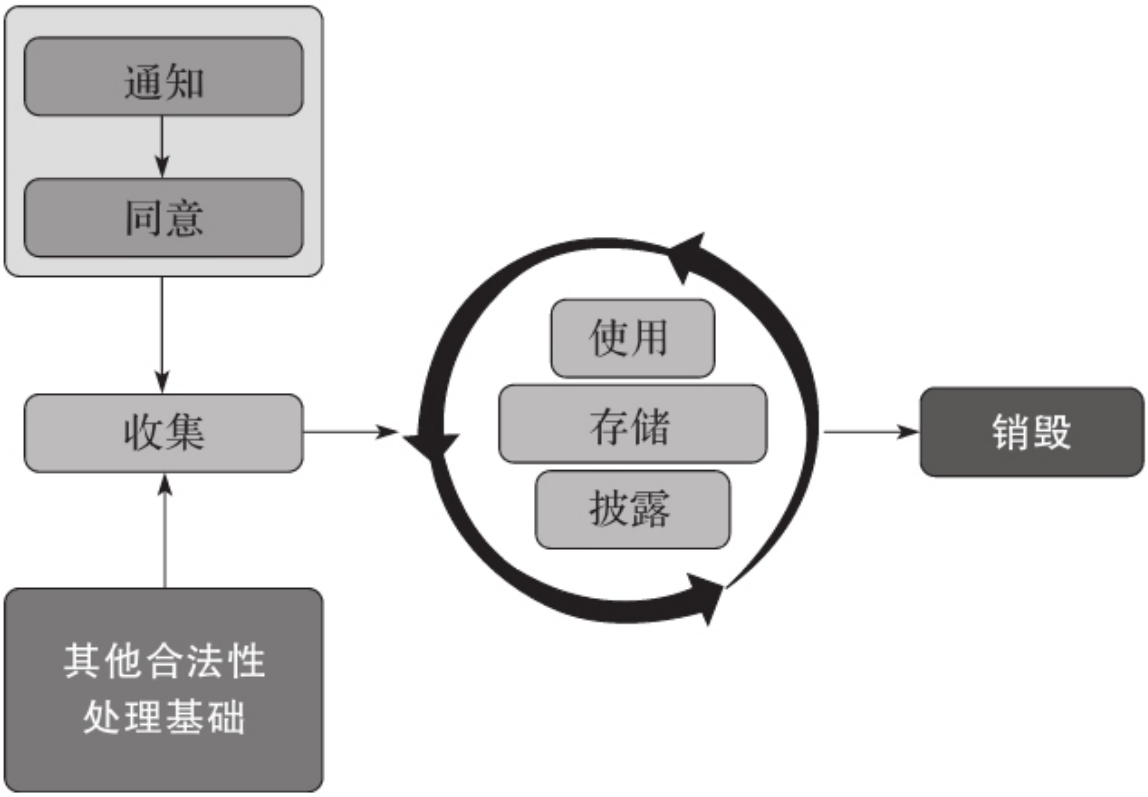

图1-2　数据全生命周期

1. 数据收集

数据收集是指获得个人信息的控制权的行为，包括由个人信息主体主动提供，通过与个人信息主体交互或记录个人信息主体行为等自动采集，以及通过共享、转让、收集公开信息等间接获取个人信息等行为。如果产品或服务的提供者提供工具供个人信息主体使用，提供者不对个人信息进行访问，则不属于本标准所称的收集。例如，离线导航软件在终端获取个人信息主体位置信息后，如果不回传至软件提供者，则不属于个人信息主体位置信息的收集[2]。需要提示的是，如果产品或服务提供者提供工具，虽然个人信息不会上传到服务器，但是在本地能访问、处理，即享有控制权，则

仍然属于收集个人信息。例如，部分App读取用户移动设备剪贴板的内容，构成数据收集。

数据收集从来源可以分为自主收集和从第三方间接获取两种。企业自主收集个人信息的，如在服务中要求用户提供账号名和密码用以创建账号，则应满足知情同意、合法、必要、正当等原则。若企业从第三方间接获取，如广告主收到广告发布平台收集的用户浏览点击广告行为，则应进一步了解第三方收集数据的合法性以及用户授权范围等。

数据收集从用户感知程度可以分为积极收集和消极收集两种。积极收集是用户可感知的，如用户自主提交的账号、个人信息档案等。消极收集是指用户相对无感的收集，如产品或服务自动采集的GPS位置、人脸信息等。例如，2021年3·15晚会上曝光某些线下门店使用具备人脸识别功能的摄像头，准确掌握用户到店情况、浏览商品情况，以及性别与年龄等，用以精准推销。用户对这种消极收集的感知更弱，数据收集中更应该履行充分告知义务。

2.数据使用

数据使用没有具体的定义，一般概括为除了存储和销毁以外的其他处理活动。这类活动一般都是基于收集时的目的或功能，若超出该目的，则需要另行告知用户且经过用户同意。数据使用的处理活动可能根据其活动的特性有所不同，如数据清洗、数据建模、数据分析等，从个人信息保护角度需要特殊关注的主要是用户画像、个性化展示以及数据汇聚。在数据使用过程中，应采取适当且必要的安全保障技术和组织措施。

3.数据披露

数据披露并非是严谨的法律概念，而是从技术逻辑角度描述将数据提供给特定和不特定的第三方，可能包括共享、转让和公开披露。

共享是个人信息控制者向其他控制者提供个人信息，且双方分别对个人信息拥有独立控制权的过程[3]。例如，支付服务提供商将订单支付结果提供给电子商务平台，而支付结果在支付服务提供商和电子商务平台有其各自的目的，因此也拥有独立控制权。

转让是将个人信息控制权由一个控制者向另一个控制者转移的过程[4]。例如，A企业独立运营的产品收集了个人信息，后来A企业被B企业并购，该产品的运营方变成B企业，则收集的这些个人信息从A企业转让给了B企业。

公开披露是指向社会或不特定人群发布信息的行为[5]。例如，用户主动在社交媒体上公布其个人相关信息，包括手机号码、家庭住址等。

数据披露环节从风险防控的角度理解，增加了参与数据处理活动的主体，个人信息保护风险可能有所增加。因此要求企业在发生数据披露时应当充分履行告知义务，包括接收个人信息的第三方名称或类型、披露个人信息的类型、披露个人信息的目的

等，并经过用户同意。而且企业应当与第三方之间定义相互的义务和因个人信息保护而承担的责任等。

4.数据存储

数据存储一般是指将个人信息存储到一定的介质上。数据存储环节主要关注的是存储期限。个人信息存储期限应当根据数据收集的目的、法律规定以及用户意愿等确定，该存储期限或存储期限的确定规则应当告知用户。

同时数据存储环节也会关注存储所采用的技术措施，如是否进行了去标识化，是否针对敏感个人信息予以加密等。

5.数据销毁

法律规定和国家标准中可能会使用“数据删除”，但是“数据销毁”一词相对来说更准确，它明确指向个人信息的彻底删除，而非仅是加密或者简单覆写。因为后两种情况可能会存在数据恢复后造成个人信息泄露的风险。在数据存储期限届满或用户行使注销账号等权利时，应当进行个人信息销毁。

[1] Travis D. Breaux, An Introduction to Privacy for Technology Professionals, 2020。
[2] 《信息安全技术 个人信息安全规范》（GB/T 35273—2020）第3.5条。
[3] 《信息安全技术 个人信息安全规范》（GB/T 35273—2020）第3.13条。
[4] 《信息安全技术 个人信息安全规范》（GB/T 35273—2020）第3.12条。
[5] 《信息安全技术 个人信息安全规范》（GB/T 35273—2020）第3.11条。