

# 第5章

## 数据全生命周期管理的目的和意义

## 5.1 数据的价值体现

对数据进行采集、传输、存储、交换和共享、处理、销毁全过程全生命周期管理,并建立完善的数据标准、数据质量、数据管理及治理、数据安全体系,其目的是充分发掘大数据所隐藏的价值,能够让数据在各行各业充分流通和融合,同时在此过程中能够进行全程的管控和溯源,了解和掌握数据的流转路径,建立数据的价值回溯和价值分配体系,让数据成为各企业及机构的核心资产,并在全产业、全社会、全球的范围有效地进行数据要素的配置、交换和交易,以最大化地发挥数据生产力的作用,促进全球数字经济的创新和发展。

数据的价值发挥体现在数据可以帮助企业和机构全面掌握生产、经营、营销、销售、供应链、服务等整体及各个环节的状况,从中发现问题及风险,找到改进和提升的关键点,对未来的产业发展、技术导向、市场及销售进行有效的预测,并用于辅助和指导企业的关键战略及决策,从而达到提效降本、保持竞争力和创新力、服务好市场和客户的目的。

数据具备很多特性,其中具有代表性的特性是数据能够发挥价值,以及驱动创新的根本原因。首先,数据在其对全局的描述、决策和预测方面,具备宏观性、中观性、微观性不同层级的掌握程度。其次,数据在其价值体现的效用前提方面,具备活性和流动性的特点。再次,数据在其融合和集成的方面,具备聚变及黑洞的效应。最后,数据在其价值发挥和行业促进方面,具备裂变和催化的作用。

### 5.1.1 数据的宏观性、中观性、微观性

数据可以帮助企业及机构对其外在及内部的环境与状况进行综合评估和掌握。根据其程度的不同,数据既具备宏观性(也称望远镜特性),又具备中观性(也称放大镜特性)和微观性(也称显微镜特性)。

宏观性指的是大数据收集的是全样本的历史数据,是所有过去数据的集大成者,依据这些数据,可以预测未来。用数据进行预测,可以预测今后的经济走势、市场形势、消费趋势、发展方向等。

Google 的 FluTrend——流感趋势预测,可以基于人们在 Google 搜索引擎上的搜索关键词,将与流感相关的关键词关联起来,比美国国家卫生署提前一周到半个月准确预测流感疫情的暴发。美国罗切斯特大学的学者和微软公司的研究者一起合作,分析了从 700 多人及其车辆上收集的超过 32 000 天的 GPS 数据,从数据中寻找模式并计算一个人某个时间会在某个地方的概率。根据他们的模型,能够预测一个人在未来 80 周的行踪,并且预测的准确率达到了 80%。也就是说,根据这些人的历史出行记录,可以预测他们未来一年半中所在的位置,这就是数据的望远镜特性,可以准确地预见未来。

数据的中观性是指可以根据从企业或机构生产经营各个来源所收集的数据,对企业或机构的现状进行一个完整的画像,帮助企业对现状有更加清晰的认知和把握,相当于用放大镜帮助他们从总体上对自身或环境、市场进行一个全面的观察和认知,梳理已有的数据资产,从中能够发现一些规律和问题,能够更好地发挥数据的价值,改进流程和环节,把控和降低风险。现在各个行业和企业所积极部署和建立的数据中台、数据大屏、数据画像,大部分的作用

用是建立一个中观的视图。如市场总量,企业的总收入,各个产品、销售部门的市场占比,逐年数据的对比等这些可视化的方法都是在中观层面上的汇总和描述。

数据的微观性指的是通过采集和汇总各个支干的数据,可以精确掌握企业或个人的最细微的细节,就像显微镜一样能够看到更加清晰、更加细致、更加微观的层面,然后就可以通过数据来做精准的运行、监控、画像及服务。如通过物联网传感器,就可以采集到每一个设备每分每秒的实时运行数据,实时监测到已发生或有可能发生的故障,并采取及时的行动来避免或减小风险和损失。

美国的电商网站亚马逊,用户 80% 的再次购买行为都是基于系统的推荐,这是因为系统记录了用户的基本信息以及他们每次的消费信息,包括家庭购物信息,这样就可以准确地掌握他们的行为、兴趣、意图和爱好,从而推测他们会喜欢什么样的商品,为他们提供精准的推荐。数据就像显微镜一样,观察到了用户最细致的信息,了解他们的一举一动。当然,这里面也涉及用户的隐私,数据有可能比用户还更了解他们自己,因此在精准服务和隐私保护两方面要做好平衡。现在有很多互联网的 App 都在非法窃取用户的隐私,并出现了利用大数据进行“杀熟”的行为,也就是利用用户的消费习惯更多地收费和牟利。因此,数据的显微镜特性也是一把双刃剑,包括我国和很多国家已经在加强这方面的数据保护及隐私立法。

### 1. 数据的活性和流动性

数据具备很多很好的特性,但是如果只有数据,不把它们很好地流通和利用起来,或者数据不能及时地采集和更新,数据的价值也得不到很好的发挥。所以还要关注数据的活性和流动性。现在社会和企业的数据已经非常庞大,尤其是一些传统企业,如医疗、旅游、金融、交通等行业,累积了几十年的数据,但这些数据有的是

纸质的放在文件柜里,有的是放在计算机里只是用来形成报表,做最基础的统计分析,数据并没有被利用起来,它们处在沉睡的状态。数据需要活动起来,把数据激活起来、唤醒起来,才能发挥其巨大的威力。数据也和资金一样,需要周转起来,流动起来,发挥其流动性。做生意的都知道,资金周转越快,周转的次数越多,就越能赚钱。数据也是一样,需要更快、更多次数地使用数据、交换数据、融合数据,才能更多地发挥它的价值。

## 2. 数据的聚变效应和黑洞效应

当把来自各种数据源的数据进行不断的聚集和融合时,就可以产生密度更大、质量更大的数据粒子,其所蕴含的价值和能量也就越大。这个聚合的过程类似于一个核聚变的过程,最后能释放出来巨大的能量。全行业的、全国性的、全球性的数据聚合起来可以爆发相当于核能量的驱动力和创新力。所以说数据是数字经济时代创新的原动力,是核动力引擎。

关于数据的黑洞效应,我们知道一个大质量的星体不停地旋转,就能形成强大的吸附力,把周边的物质都吸收进去,甚至连光线都不能逃逸,最终形成一个黑洞。如果整合多行业、多源的数据,发挥其活性和流动性,数据的质量越来越大,数据流转速度越来越快,就可以把周边所有相关的数据、资源、人才等都全部吸附过去,形成一个巨大的数据黑洞,最终只要跟这个数据黑洞发生交集的都会被“吞噬”进去。

我们预测未来全球就像我们的宇宙一样,可以形成多个数据黑洞,现在的很多互联网大平台,由于其本身累积了大量的数据,同时又在不停地收集和整合行业数据,已经形成了一些数据黑洞。从其积极意义来看,数据越来越集中,价值流转的速度越来越快,极大地发挥了数据的价值。但是,由于数据过于集中,也产生了数据垄断,数据定价权被掌握,对于行业的创新和开放式发展也有不

利的因素。

### 3. 数据的裂变效应和催化效应

数据通过聚变累积了巨大的价值和能量,当把所聚合的数据通过分析加工应用到各行各业中去时,就发生了数据的裂变。如用户行为数据,几十亿用户的行为数据,可以应用到跟用户相关的医疗、教育、社交、休闲、娱乐、金融等各个产业方向,一份数据就有了不同行业的应用裂变,其所产生的数据及应用价值是不可计量的。

越来越多的人意识到数据所蕴含的巨大价值。数据被誉为新时代的黄金和石油,然而数据有一个特性是黄金和石油不能比拟的,使得它比黄金和石油都更有价值。这就是数据的催化剂特性。催化剂可以加速化学反应的过程,但它本身并不损耗。同样,数据在使用过程中可以加速整个生产、经营和商业营销、销售、服务的过程,但数据本身并不损耗,怎么用数据都在那里。数据可以重复使用,而且数据还是越用越多,在使用的过程中又产生了新的数据,因而可能越用价值越高。数据跟多种数据源交叉使用时,价值沉淀就越来越大。同样,数据可以深入应用到全行业,可以循环使用。任何一个其他的行业要素都是会损耗的,用完就没有了,但数据就可以一直使用,加速和催化所有的产业和环节,还可能越用越值钱。数据可以说是最值钱的生产要素,这是其他生产资料不可比拟的。

举例来说,美国的政府数据开放网站 [www.data.gov](http://www.data.gov) 累积了美国几十个行业和部委数十年的数据,而且全面开放给社会,在其上还建立了开放共享的工具集和应用,这样就使得数据和工具相互促进,发挥了聚合和催化的作用,其一年带动的创新产值是数万亿美金。其中一个例子是美国的 Climate 创业公司,基于上述政府开放数据网站,汇总了 250 万个地点的气象测量数据和各个主要

气候模型的天气预报,同时综合 1500 亿个土壤观测记录,对这些数据进行处理,生成出 10 万亿个天气模拟数据点,为农业生产提供保险服务。Climate 公司的几位联合创始人是 Google 的早期员工,他们为天气保险的投保人开发了一种自助式互联网服务,此前这类保险只能通过定制的方式进行柜台交易。现在,客户可以登录 Climate 公司的网站,确定特定时间段内需要投保的气温和/或降水量范围。平台收到订单后,就会在 100ms 内综合分析天气预报、近 30 年来的国家气象局数据,以及用户所在地的地质调查数据,并根据气候变化,对分析结果进行微调。得出结果后,就会作为保险商,给用户开出保费。投保人如果因为意外天气而受到损失,就能自动获得赔偿。Climate 公司创立仅 3 年就被美国最大的农业公司孟山都以近 10 亿美元收购。类似这样的数据创新公司在数据开放平台上比比皆是。近些年来,中国各个城市也都在建立数据共享和开放平台,用以促进数据的开放和创新。

## 5.1.2 数据可视化

### 1. 数据可视化概述

数据可视化是借助于图形化手段,清晰、有效地传达与沟通信息。数据可视化不一定因为要实现其功能用途而过于理论化,或是单调枯燥,也不需要为了绚丽多彩而显得极端复杂。为了有效地传达信息及概念,美学形式与功能需要齐头并进,通过直观地传达关键的方面与特征,从而实现对于相当稀疏而又复杂的数据集的深入洞察。

数据可视化与信息图形、信息可视化、科学可视化以及统计图形密切相关。当前,在研究、教学和开发领域,数据可视化是一个很活跃而又关键的领域。尤其是最近元宇宙概念的兴起,未来人

类可能更多地在数字空间进行交互,那么关于数据及信息的三维甚至更多维的表达也是至关重要的。

## 2. 为何进行数据可视化

数据可视化领域的起源可以追溯到 20 世纪 50 年代计算机图形学的早期。当时,人们利用计算机创建出了首批图形图表。1987 年,由布鲁斯·麦考梅克、汤姆斯·蒂凡提和玛克辛·布朗所编写的美国国家科学基金会报告《科学计算之中的可视化》,强调了新的基于计算机的可视化技术方法的必要性。随着计算机运算能力的迅速提升,人们创建了规模越来越大、复杂程度越来越高的数值模型,从而造就了形形色色体积庞大的数值型数据集。同时,人们不但利用医学扫描仪和显微镜之类的数据采集设备产生大型的数据集,而且还利用可以保存文本、数值和多媒体信息的大型数据库来收集数据。因而,就需要高级的计算机图形学技术与方法来处理和可视化这些规模庞大的数据集。

早期的可视化主要是在科学与工程实践当中对计算机建模和模拟的运用。后来,可视化也日益尤为关注数据,包括那些来自商业、财经、行政管理、数字媒体等方面的大型异构数据集合。当今,数据可视化指的是采用较为高级的技术方法,允许利用图形、图像处理、计算机视觉和用户界面,通过表达、建模及对立体、表面、属性和动画的显示,对数据加以可视化解释。与立体建模之类的特殊技术方法相比,数据可视化所涵盖的技术方法要广泛得多。

先进的数据可视化技术和方法不再局限于通过关系数据表来观察和分析数据信息,还能以更直观的方式看到数据及其结构关系。大量的数据集构成较为丰富和立体的数据图像,同时将数据的各个属性值以多维数据的形式表示,可以从不同的维度观察数据,从而对数据进行更深入的观察和分析。



### 3. 数据可视化交互形式

从数据展示的角度来看,可视化技术主要是针对数据的结构、功能、关联关系、发展趋势等几方面进行展示。传统的可视化展示方式一般是图表、图形、数字化模型及三维展示、虚拟现实/增强现实展示等。随着大数据的兴起与发展,互联网、社交网络、地理信息系统、企业商业智能、社会公共服务等主流应用领域逐渐催生了几种特征鲜明的信息类型,主要包括文本、网络图、时空及多维数据等。这些与大数据密切相关的信息类型与多维数据模型交叉融合,将成为数据可视化的主要研究和发展领域,因此也衍生了很多新形式的展现和交互模式。下面介绍6种数据可视化及互动技术:二维展示技术、三维渲染技术、虚拟现实技术、增强现实技术、可穿戴技术和可植入设备技术。

#### 1) 二维展示技术

二维展示技术包括标准图表(柱状图、折线图、饼图等)、时间序列(Time Series)、层级树状图(Hierarchical Tree Map)、时间轴、地图、网络图、信息图等。

近几年涌现出了一大批基于二维展示技术的数据可视化服务公司。以 Google 为代表的几家公司提供的可视化服务尤其突出。Google 的 Charts 提供了用户在网页上以图形方式展示数据的接口,Charts 支持饼图、折线图、柱状图、区域填充图、散点图、维恩图、仪表盘等多种形式,并可以设置图中各部分的颜色、形状、间隔等细节。还有很多其他开源的工具支持文本可视化、地图、社交图谱等各种数据的可视化呈现。图 5.1 是一个社交图谱可视化的例子。

#### 2) 三维渲染技术

三维渲染技术是近年来发展迅速和备受关注的行业,在数字娱乐、虚拟现实、工业设计、实时仿真、数字城市等各个领域都有着

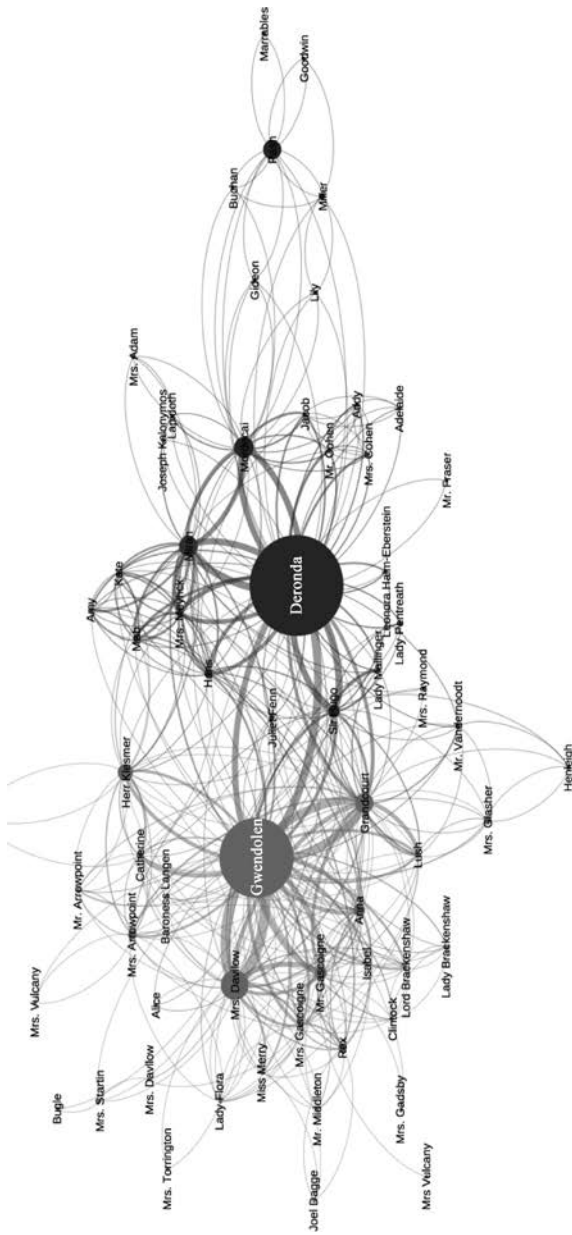


图 5.1 社交网络关系图谱

十分广泛的应用。在数字娱乐领域,提到三维动画渲染,人们马上就会联想到皮克斯公司。皮克斯是一家专门制作计算机动画的公司,其制作的《怪兽公司》《虫虫危机》《海底总动员》《料理鼠王》等动画电影系列,都受到全球观众的热捧。三维特技在科幻影片《阿凡达》中更是被发挥得淋漓尽致。该片也因为震撼的特技效果而获得了全球电影史上的最高票房,并获得了第 82 届奥斯卡最佳艺术指导、最佳摄影和最佳特效 3 项奖项,以及第 67 届金球奖最佳导演奖和最佳影片奖。

在工业设计领域,目前在建筑、飞机、轮船、汽车、机床等设备的设计中已经普遍用到三维技术,它使设计师可以在屏幕上随时变更设计方案,进行快速验证。在当今的数字城市、智慧城市建设中,三维技术也展现了巨大的能量,它不仅能够模拟整个城市、园区、建筑、室内的建设效果,还能结合控制参数,实时仿真出动态反应场景,如变电站的控制、水库的监测等。采用三维技术,可以大大节省实际生产和制造的时间和成本,同时直观地展示出最终的效果,能够高效地进行互动调整等。

### 3) 虚拟现实技术

虚拟现实(Virtual Reality,VR)技术是由美国 VPL 公司创始人拉尼尔(Jaron Lanier)在 20 世纪 80 年代初提出的,也称灵境技术或人工环境。作为一项尖端科技,虚拟现实集成了计算机图形技术、计算机仿真技术、人工智能、传感技术、显示技术、网络并行处理等技术的最新发展成果,是一种由计算机生成的高技术模拟系统。

这种技术的特点在于计算机产生一种人为虚拟的环境,这种虚拟的环境是通过计算机图形构成的三维数字模型,并编制到计算机中去生成一个以视觉感受为主,同时包括听觉、触觉的综合可感知的人工环境,从而使人产生一种沉浸于这个环境的感觉,人可以直接观察、操作、触摸、检测周围环境及事物的内在变化,并能与

之发生“交互”作用,使人和计算机很好地“融为一体”,给人一种“身临其境”的感觉。

一般的虚拟现实系统主要由专业图形处理计算机、应用软件系统、输入设备和演示设备等组成。虚拟现实技术的特征之一就是人机之间的交互性。为了实现人机之间充分交换信息,必须设计特殊输入工具和演示设备,以识别人的各种输入命令,且提供相应反馈信息,实现真正的仿真效果。不同的项目可以根据实际应用选择使用不同的工具,主要包括头盔式显示器、跟踪器、传感手套、屏幕式或房式立体显示系统、三维立体声音生成装置等。

#### 4) 增强现实技术

增强现实(Augmented Reality, AR)技术是在虚拟现实技术的基础上发展起来的,它是把计算机产生的虚拟物体和场景叠加到现实场景中,用这种混合的模式增强用户对场景的感知。它是一种全新的人机交互技术,利用这样一种技术,可以模拟真实的现场景观,它是以交互性和构想为基本特征的计算机高级人机界面。使用者不仅能够通过增强现实系统感受到在客观物理世界中所经历的“身临其境”的逼真性,而且能够突破空间、时间以及其他客观限制,感受到在真实世界中无法亲身经历的体验。

AR技术包含了图形图像学、可视化技术、实时交互技术、多传感器融合等新技术和新手段,系统具有3个突出的特点:真实世界和虚拟世界的信息集成、实时交互性、在三维尺度空间中增添定位虚拟物体。AR技术可以广泛应用于军事、医疗、古迹保护、建筑、教育、工程、影视、娱乐等领域。在军事领域,用于尖端武器、飞行器的研发、虚拟训练等;在医疗领域,用于帮助医生进行手术部位的精确定位;在古迹复原和数字化文化遗产保护领域,将文化古迹的信息以增强现实的方式提供给参观者,使参观者不仅可以通过专用的显示器(头盔式、眼镜式等)看到古迹的文字解说,还能看到遗址上残缺部分的虚拟重构;在影视领域,通过AR技术可以在转

播体育比赛时实时地将辅助信息叠加到画面中；在商品展示中，可以将家具和电器叠加到顾客的客厅中查看实时效果等。

#### 5) 可穿戴技术

可穿戴技术主要是探索和创造能直接穿在身上，或者整合进用户的衣服或配饰的设备的科学技术，这种技术互动实现了可视化与实时的有机集成。可穿戴技术是20世纪60年代美国麻省理工学院媒体实验室提出的创新技术，利用该技术可以把多媒体、传感器和无线通信等技术嵌入人们的衣着中，可支持手势和眼动操作等多种交互方式。其目的是通过“内在连通性”实现快速的数据获取，通过超快的分享内容能力高效地保持社交联系，摆脱传统的手持设备而获得无缝的网络访问体验。

20世纪60年代，可穿戴技术逐渐兴起；到70年代，发明家Alan Lewis打造的配有数码相机功能的可穿戴式计算机能预测赌场轮盘的结果。1977年，Smith-Kettlewell研究所视觉科学院的C. C. Colin为盲人做了一款背心，把头戴式摄像头获得的图像通过背心上的网格转换为触觉意象，让盲人也能“看”得见。自2012年4月Google宣布其Google Project Glass的未来眼镜研发项目后，各大科技公司纷纷在可穿戴技术应用上加大研发力度。2013年，苹果公司密集曝光了其智能手表Apple Watch的一系列新功能并于2014年进行了产品发布；索尼公司于2013年8月底推出了Smart Watch的第二代产品；三星公司则在2013年9月推出了智能手表产品Galaxy Gear。随着移动通信、图像技术、人工智能等技术的不断发展及创新融合，在全球应用和体验式消费的驱动下，可穿戴设备迅速发展，已成为全球增长极快的高科技市场之一。据统计，全球可穿戴设备出货量从2014年的0.29亿部增长至2021年的5.34亿部，预计到2024年将达到6.37亿部。

可穿戴技术是近几年科技界热门的趋势之一。在每次大型科技盛会上，面向个人消费者的可穿戴设备数量都呈现指数级增长，

从智能手环、智能手表、智能手套到智能眼镜、智能头盔等。可穿戴设备不仅仅是硬件设备,更是可以通过软件支持及数据交互、云端交互来实现强大功能拓展的设备。可穿戴设备具有快速的信息抓取、处理和查询能力,以及更准确的判断决策能力,在这种设备的帮助下,人们的行为模式和行动效率也将得到改善和提高。

#### 6) 可植入设备技术

由于医学发展的需要,为修复受损功能和维持身体运作,人们发明了医疗设备植入人体的技术,如心脏起搏器和耳蜗的植入。随着医学植入技术和无线传感技术的发展,越来越多的科技人员开始研究可植入人体的高科技应用设备,以便通过植入设备来监测、管理和改善人的身体状况,让身体运作得更加协调有序或者实现定位追踪等特殊目的。随着应用场景的不断发展及科技的突破,以埃隆·马斯克为代表的一些人已经在探索脑机接口(Brain Computer Interface,BCI),即实现人脑与计算机的互联。早期布朗大学的 BrainGate 团队,深入研究了如何实现人的大脑与计算机对接。初步研究显示,在人脑中植入婴儿版阿司匹林大小的电极,神经信号可被计算机实时解码,并用于操控外部设备。而现如今,脑机接口离商用化也就一步之遥。脑机接口的落地场景主要在医疗、娱乐、教育等领域,具体如下。

在医疗领域,用于运动控制的脑机接口系统记录与思维、感知和运动意图相关的神经活动,将脑信号解码为用于输出设备的命令,并通过输出设备执行用户的预期动作。用于感觉增强的脑机接口系统将环境刺激转换为中枢神经系统可解释的神经信号。这两种类型的系统都有可能通过促进用户与环境的交互来减少残疾。脑机接口技术被用于康复环境中,既可作为神经假体来替代失去的功能,也可作为旨在加速神经恢复的潜在可塑性增强治疗工具。受益于运动和体感脑机接口系统的人群包括脊髓损伤、运动神经元疾病、肢体截肢和中风患者。

在娱乐领域,当脑机接口技术发展成熟后,使用者将无须佩戴任何 AR 或 VR 等体外成像设备,而是在脑部植入一颗芯片,通过模拟神经信号接入虚拟世界,就可以获得逼真的体验。此外,脑机接口或将成为下一代人机交互技术的主要形式,也将成为元宇宙的终极接入方式。

可植入设备的研究已经不仅仅局限于医疗领域,但是由于可植入设备的电极、传感器等材料可能被身体组织吸附产生的副作用,医疗领域的皮下植入技术规范能否在可植入设备大量普及的情况下得到严格执行,设备能否长时间供电等技术问题,此外还有大量的伦理道德问题都有待研究和探讨,因此可植入设备的未来发展还有待时日。

### 5.1.3 数据可分析

数据的各种价值特性,以及对数据的全生命周期管理的目的都是充分发掘数据的价值,发挥数据的生产力作用。而其前提是能够对数据进行加工、分析和处理,从中发掘出有价值的信息,并用于生产、生活、工作、经营的实际活动中,体现出其价值和效应。

大数据正成为新一代信息技术融合应用的核心,为云计算、物联网、移动互联网、人工智能等各项新一代信息技术相关的应用提供坚实的支撑,数据分析是这一切的基础。同时,大数据蕴含着巨大的社会、经济和商业价值。数据分析会催生一大批面向大数据市场的新模式、新技术、新产品和新服务,进而促进信息产业的加速增长。此外,各行业对大数据的实际需求能够孵化和衍生出一大批新技术和新产品,来解决面临的大数据问题,促进科技创新。对于数据的深度利用,将能够从数据中挖掘出潜在的应用需求、商业模式、管理模式和服务模式,这些模式的应用将成为开发新产品和新服务的驱动力。

### 1. 数据分析对于企业的意义

数据分析对于企业具有多重意义。企业通过对自身累积的数据和外部收集的数据进行分析,可以实现数据的深度挖掘和利用,形成智能决策,在企业运营中提高效率,节省成本;在市场竞争中制定正确的市场战略,把握市场先机,规避市场风险;在市场营销中全面掌握用户需求,进行精准营销和个性化服务。例如,零售企业可通过对数据的实时分析掌握市场动态并迅速采取应对措施,通过精准营销增加营业收入;工业制造企业可通过整合来自研发、工程和制造部门的数据,实行并行工程,缩短产品上市时间并提高质量。各类企业还可从产品开发、生产和销售的历史大数据中找到创新的源泉,从客户和消费者的大数据中寻找新的合作伙伴,以及从售后反馈的大数据中发现额外的增值服务,从而改善现有产品和服务,创新业务模式。

在数据时代,企业的决策正在从“应用驱动”转向“数据驱动”,能够有效分析利用数据并将其转换为生产力的企业,将具备核心竞争力,成为行业领导者。

#### 1) 智能决策

数据分析对于企业的价值和意义,首先体现在宏观的战略层面,通过对政策、行业、市场、产品及技术在整体上的分析,可以综合把握国内外这些方面的发展态势,制定正确的企业战略,应对不确定的变化,综合把控和降低风险。如在华为公司的发展早期,就是敏感地把握到了数字程控交换机将成为通信行业的主要发展方向,因而在战略上进行资金和人才的全部投入,造就了华为公司在市场上的领先地位。类似这样的战略决策现今可以通过数据分析来实现。

在中观层面,通过对企业生产经营数据的实时分析,可以及时地掌握市场及销售动态,发现问题和机会,并能够迅速进行决策和



响应,如对热销产品的供应链及资金的配套,进行及时的补充,就能够避免因为货源的储备不足而丧失市场机会。

在微观层面,则能够根据市场及用户的反馈,实时地对产品细节及营销策略做出微调,以满足不同用户的需求,提升服务质量。世界知名的快时尚品牌 ZARA,会每天收集社交网络信息,以及门店的消费者的反馈,根据汇总的数据微调某一品类的衣服的样式、尺寸,来更适应用户的个性化需求,这样的微观智能决策也帮助 ZARA 成为时尚界的翘楚。

### 2) 增加收入

企业经营的最直接目标还是增加收入。增加收入有多种办法,如提升产品的市场占有率、提升产品的品牌价值、扩大生产的产量、提高产品的销量等。而这些都不能够通过拍脑袋来决定,都需要通过数据分析来进行智能定位的决策。例如在产品的品牌塑造方面,就需要精准地定位产品的目标人群,包括他们的价值取向和品牌认同度、他们的消费倾向等。而在产品的市场方面,则需要进行竞品分析、热点分析、时尚流行分析等。在产品的销售方面,则需要进行渠道、销量、片区、部门、季节等多维度的分析对比,才能达到提效增收的目的。

### 3) 提高利润

采取前述的办法在提高产品知名度、品牌、市场占有率和销量的基础上,也可以实现高额的利润。但增加收入并不等同于利润的提升,因为还涉及成本的投入和计算,以及经营和管理的效率。提高利润的主要方法在于提效降本。数据分析在企业的供应链、生产、管理、经营、销售等各个环节都能够进行细致化的分析和监控,从而能够在总体上和各个环节上实现提效降本,从而增加利润。提高效率的办法是实现各个合作企业、各个部门,以及各个环节的协同效率,减少摩擦和损失。降低成本则是在办公设施、管理成本、销售成本、材料成本、生产成本等各方面避免浪费,

节约支出。通过数据分析的中观及微观效应,可以细致入微地发现问题,解决问题,规避风险,最终提高企业的利润。

#### 4) 提升客户满意度和忠诚度

提升客户满意度和忠诚度在于更好地了解和掌握客户的需求及喜好,并做好相应的品牌定位及服务。这需要通过数据收集及数据分析,来定位用户群,分析用户内容偏好,分析用户行为偏好,建立受众分群模型,制定渠道和创意策略,完成投放评估效果等,不断把控营销与服务的质量与效果,并结合历史数据分析实现从效果监测转向效果预测。

在具体操作层面,一方面需要从各个渠道采集相关的数据,尤其是随着电商平台、互联网的应用以及社交媒体的发展,每一个潜在客户都在网络上留下了很多的数据,通过对海量数据的多维度重组,使得企业能够通过精准细致的数据引导自身营销策略的改变。另外,企业之间也在推动各平台间的内容、用户、广告投放的全面打通和共享,通过用户关系链的融合以及网络媒体的社会化重组,给企业带来更好的精准营销和服务效果。数据的采集也不完全局限于互联网的方式,社会调查、门店及渠道等也可以作为有效的方式。另一方面,数据分析的结果也需要通过适当的渠道和方式再传递给客户,这同样可以通过线上和线下的方式,建立好全渠道的用户营销和服务。

## 2. 追溯追责

通过数据的全生命周期管理,可以建立完整的数据闭环,涉及全行业、全环节的数据及其对应的产生、收集及使用的过程,这样就可以形成完整的追溯过程,而在任意一个环节所发生的问题、失误及损失,都能找到其对应的责任主体,从而也实现追责的功能。

下面举例来说明,如要建立食品追溯体系,首先需要的就是通

过信息化手段覆盖从种植养殖、加工、包装、物流、销售等生产及经营过程,完整地记录各环节的经营流通数据,以便后续可追溯。食品安全追溯体系还需全面采集和记录食品的生产、仓储、分销、物流运输、市场巡检及消费者等信息,以及产品名称、执行标准、配料、生产工艺、标签标识等数据,并进行跟踪和分析。这些需要建立在开放性、集成性良好的追溯追责平台上,才能真正实现其目的。

### 3. 监控和监督

监控和监督与上述的追溯追责有一定的相似性。在其基础上,还需要在行政地域、管理部门之间实现信息的跨区域管理和共享,避免由于现有行政区域的划分,各地政府职能管理部门形成地方保护,监管部门出现相互扯皮、推诿等不良现象,还有监管部门条块分割造成的监管信息分散、信息内容单一、部门之间信息沟通不畅等问题。

对数据进行全生命周期的综合采集、应用和管理,可优化职能监管部门的资源配置,制定出较好的统筹与协调解决方案,达到有效的行业监控监督。同时,在合理配置监管部门职能的同时,将极大提高信息的利用效率,节约成本,发挥数据的协同作用。

### 4. 洞察探寻

对数据的洞察探寻在于对数据的深度理解,对数据的各个维度进行深度的分析,了解其背后深层次的原因和导向。这需要结合数据分析及数据的可视化,进行交互式的分析和探索。例如,一款产品销量的突然下滑,一个季度的业绩的突然暴跌,都需要到数据里面去找到其根本原因,并能够找到解决的办法,它可能是营销策略、产品定价的失误,也可能是相关的社会事件,还可能是关键人物的离职,这些都是数据洞察能够给出的答案。

## 5. 挖掘商机

数据分析能够有效地发现市场机会,找到消费品之间的关联,了解到用户的消费习惯,这些都是可以利用的商业机会。如数据分析可以发现市场热销商品,那么及时地进行相应产品的备货和销售,就能马上赚取利润。同样,美国的沃尔玛是最早利用“购物篮关联分析方法”的大型零售商,发现了“啤酒与尿布”的关联,因而把两样商品放置在一块,提高了商品的销量。同样发现了在飓风季节,客户对甜品需求的上升,因而把甜品与手电筒、雨具放置在相邻的位置,来进行促销,都取得了良好的效果。数据同样还可以帮助新产品的研发,小米公司通过建立互联网的产品沟通渠道,与上百万人建立联系,让他们参与手机产品的设计及研发,才取得了第一款小米手机的成功。

## 6. 预测未来

预测未来是基于数据的望远镜特性。基于历史数据可以建立对未来的预测模型,从而可以有效地分析消费潮流、用户接受度、产品销量、疾病传播概率、故障可能发生的时间等。大型互联网公司在对用户页面、产品页面做出大改动之前,都会应用分析预测模型来预测用户对新网页的接受度,哪些用户可能流失,流失率会是多少等,并会在页面切换之前,采用 A/B 测试的方式,将一部分用户引流到新网页,来验证相关的预测是否准确。数据分析甚至被用来预测选民的支持程度,帮助美国总统进行竞选策略的制定,并成功地帮助包括奥巴马、特朗普等登上总统宝座。

### 5.1.4 有的放矢地分析数据

对于企业的数据分析业务来说,进行全方位的数据采集,建立

数据中台及算法模型,进行数据全生命周期的管理,这些都意味着相对比较大的建设成本和开支。如果没有相对清晰地分析目标和预期的效果,而是进行泛泛的探索,或是不管三七二十一,什么都想上,那么很可能就会面临效益远远小于投入的尴尬结局。企业的数据库建设可以本着循序渐进,首先解决企业的难点、痛点方面来进行,对关键数据进行收集,制定明确的数据分析目标,分析的结果可以及时地应用在企业的经营决策中,并带来相对可观的效益和效果。

### 5.1.5 不同行业数据分析的侧重点

在不同的行业应用中,对数据分析的需求及侧重点也有一定的不同。在互联网金融领域,在线的金融服务已经逐步成为主流,实体的门店职能越来越弱化。金融产品及服务可能重点关注风险的管理和控制,因此基于数据挖掘的客户识别和分类将成为风险管理的主要手段,对风险管理也需要动态、实时的分析和监测,而非事后的统计和评价。

在能源电力行业,对于工业企业的用电信息需要基于实时的电表采集的信息,进行分析和统计,用于电价的制定,以及电力资源的配置。因此基于大量数据的多维度统计和分析预测就成为电力行业的主要分析场景。同时,对电力设备的监控及故障分析预测,也是保障电力网络有效安全运行的基础。

在民航领域,由于全球民航业务的发展,以及各大航空公司之间的激烈竞争,提高效率、节约成本、拓展服务模式、提高服务质量已经成为各个航空公司的重点发展策略。因而对服务产品的定制、对用户的精准产品推荐、提升服务的质量以及效果的监测,就是民航业务的主要数据分析场景。

在教育领域,则着重要为老师和学生减负。因而精准的智能

教学辅导工具,学生的精准行为及知识点、成绩分析,以及学生的综合能力的训练与素质培养就成为教育领域数据分析的重点。这可以通过建立能力与知识模型、学生的学习习惯、解题及分析的方法等关联模型,最终达到因材施教、因人施教的个性化培养的目的。

## 5.2 数据资产

### 5.2.1 什么是数据资产

从前面的论述可以看到,数据已经不单纯是信息的载体,它是可以直接创造价值的核心资源,从这个角度来看,数据已经上升成了企业的资产。数据资产是企业拥有的,带有明确权属的,能给企业带来收益的数据资源。相比单纯的数据,数据资产有以下特征。

(1) 数据资产是具有价值的,也就是有价值的数据资源才可以是数据资产。只有当数据经过一定形式的加工、处理及利用,产生了相应的价值,才能被称为数据资产。

(2) 数据资产是需要加工的。企业内部很少一部分的数据资源可以直接作为数据资产给企业创造价值,大部分的源数据无法形成资产价值,需要经过整理、汇总、聚合、加工、处理形成新的数据资源,才能为企业创造价值。

(3) 数据资产是需要管理的。数据资源作为企业无形资产的一部分,是需要管理才能为企业创造更大的价值,而经过管理的数据资产才可能为企业创造价值最大化。

(4) 数据资产是能够应用的。数据形成资产,能够应用于企业的业务经营,助力企业发展,提升企业的资产运营能力,创造企业价值。

## 5.2.2 如何进行数据资源化

数据资源化是数据市场化发展和价值挖掘的首要问题。如今,数据的价值被越来越多的企业发现并利用,数据在国家、企业和社会层面渐渐成为重要的战略资源,成为新的战略制高点。但目前,无论是行业数据还是政务数据,开放程度极为有限,数据孤岛和数据保护现象还相当严重,这需要政府和企业进一步转变意识,加强数据的开放和共享,加大数据资源化的建设。

数据资源化是指将原始数据转换为具备一定的潜在价值的数字资源的过程,是数据资产化的必要前提。针对数据,需要有一套完善的数据资源管理流程和体系,才能完成数据到数字资源的转换。数字资源管理是以数据管理为基础,包含数据标准、数据质量、数据价值、数据共享和数据开发、数据安全等管理职能,对数据进行标准化建设及质量管理,提高数字资源价值增值能力,并保障数据安全。在企业数据资源化过程中,数据治理工作的实施至关重要。实施数据治理能够确保企业内数据的准确性、一致性、时效性和完整性,提升数据质量,保障数据安全,推动内外部数据流通。通过数据治理这道必要的闸口,企业内部的原始数据才得以转换为有一定使用价值的数字资源,为构建全面有效的、切合实际的数字资产管理体系做好准备。

在数据资源化过程中,还需要明晰数字资源主体的关系,关注如何借助数字资源为数据所有者、数据经营者和数据使用者带来数据价值。分散在各处的原始数据,经数据采集、数据加工后拥有数据所有者属性,数据转换为数字资源后更是获得交易属性,所以需要明晰数字资源主体的责权利关系。由于数据隐私及数据价值的存在,数据所有者不能简单地让渡数据所有权,因此在数字资源

化过程中必须明晰数据所有者、数据经营者和数据使用者三方的主体和权属。

数据所有者应完整占据数据资源的产权,可以自由决定是否委托数据资源经营权。数据经营者接受数据所有者委托,享有数据经营权和获取合法收益的权利。数据使用者依据数据经营者赋予的合法权益享有数据资源及其衍生物的数据价值,从而实现数据价值传递、数据价值变现。

### 5.2.3 如何进行数据资产化

数据资产化是将数据资源转换为能为企业产生价值的数据资产的过程。数据资产化是数据交易的实现途径。数据要素要想流动起来,打破数据孤岛,真正成为一种生产要素,就必须能够市场化交易或者说市场化衡量,其中资产化就是必须要完成的工作。

企业的数据资产化总体上需要经历以下5个步骤。

#### 1. 业务数据化和归一化

业务数据化是企业数字化转型的第一步,也是数据资产化的前提。所有的业务流程、业务数据都能够进行数字化的采集、存储,是建立数据资源的基础。同时,还需对多源数据进行融合和集成,同一实体的数据能够做到归一,才能为进一步的数据治理及管理做好准备。

#### 2. 建立数据标准

数据资产化需要建立数据管理、使用、交易的共识及标准,因此必须做好数据资产的质量管理和标准化建设。对企业的资产需要做好盘点、规划及架构,建立统一的数据字段标准和体系,促进



并提升数据资产标准的一致性。

### 3. 数据分类分级,做好隐私评级

数据资产的使用离不开数据的分类管理及隐私管理。可以通过设立数据资产类别,对数据资产的涉密及隐私程度进行分类,降低数据资产使用过程的涉密及隐私泄露风险。

### 4. 数据的价值挖掘

对数据进行加工、处理、挖掘,将数据标签化、价值化,让数据可持续、可应用,实现数据资源到数据资产的转换。这里面还要注重成本与效益的问题。

### 5. 数据的产品化和服务化

数据一定要能够用起来,能够形成产品、商品、服务,能够被购买、被交易、被使用,产生利润,才能真正实现数据的价值变现。

## 5.2.4 数据资产化面临的难题

全球数据资产化的进程实际上都还开展不久,很多政府及行业都还在建立基础的数据资产管理体系,以及数据资源的治理体系,在很多领域都还处于探索和实践的过程中,因此,还面临很多问题。其中有代表性的几大难题是数据资产确权、数据资产价值评估、数据资产交易和数据资产安全。

### 1. 数据资产确权

数据资产确权即数据资产的权属界定,是数据资产化的基础,也是数据交易的基础。在数据服务及交易过程中的所有参与方,需要明确各自的责、权、利。数据的权属(也称数权)不同于传统的

物权,物权是对实物的直接支配,但数权在数据的全生命周期中有不同的主体和权属,如前述的所有权、经营权、使用权等。由于数据所有权可以转移、难以界定,因此数权界定需要综合考虑主体及数据资产使用和交易的全过程。

数权可以简单理解为在法律意义上,哪个企业拥有数据的所有权、经营权和使用权等。数权需要明确相应的权利和责任,包括谁能够创建、读取、修改、复制、共享或删除哪些数据,谁享受数据的哪种收益等。即使企业对外开放或共享数据,在法律保护下,企业仍没有失去所有权、经营权和使用权。在数据资产具有可复制性的不利背景下,数权可以保障数据资产的安全和合法收益。

目前,各国在数权方面都进行了不同程度的探索,数权一般都会依据本国的相关法律和文化进行确定,如美国将个人数据置于传统隐私权的架构下,通过制定各个行业的行业隐私法,形成部门立法+行业自律的体制。2018年5月,《通用数据保护条例》(GDPR)在欧盟正式生效。通过数据立法顶层设计,加强数据主权建设,形成个人数据和非个人数据的二元架构。日本对于数据的原则是自由流通为原则,特殊保护例外。通过设立数据银行以及个人数据商店(PDS)对个人数据进行管理,在获得个人明确授意的前提下,将数据作为资产提供给数据交易市场进行开发和利用。

数权的界定需要记录、验证、跟踪、溯源数据资产的产生、使用、交易、转移的各个过程,可能需要多种包括法律、规范、管理、技术等方法 and 手段,所以也面临这几个方面的挑战。

## 2. 数据资产价值评估

数据资产价值评估是数据价值变现,即数据销售、数据交易的基础。数据资产作为一种全新的资产形态,其后又是流动的、可变的数数据,因而其价值评估方式比传统资产的价值评估更加困难、更

具挑战性,也缺乏先例的参考。首先,并非所有的数据都可视为资产,它首先是能够为企业产生价值的数据资源。数据资源转换为资产后,仍有类别和价值的不同。数据经过深度的加工、交易和流通之后,其价值的评估就更加困难。

目前,数据资产的价值估值方法主要类比无形资产,有成本法、收益法和市场法3类。成本法根据数据资产的成本构成,测算数据资产价格,与数据资产价值的重新获取或建立数据资产所需成本紧密挂钩;收益法基于目标数据资产的预期应用场景,通过预期经济效益折现,反映数据资产投入使用后的收益能力,要求数据资产收益可预测;市场法适用于市场上具备一定数量且可比的数据资产的估价对象,通过相同或相似数据资产的对比和差异因素的调整,反映目标数据资产市场价值。但由于数据可以应用于不同的场景,其来源及构成、加工方法也有很大的不同,因此在实际使用这些方法时,也都有难以适用的情况。

### 3. 数据资产交易

我国自2014年以来,就出现了一批数据交易平台和数据交易机构,包括贵阳大数据交易所、东湖大数据交易所、长江大数据交易中心、中关村大数据交易平台等。近年来又先后成立了上海数据交易所、北京数据交易所、深圳数据交易所等国家级的数据交易平台,力图完善数据交易配套制度,针对数据交易全过程提供一系列制度规范,涵盖从数据交易所、数据交易主体到数据交易生态体系各类办法、规范、指引及标准,让数据流通交易有规可循、有章可依。

数据交易产业链的参与方包括数据供应方、数据需求方、平台运营方和行业监管方。数据交易的模式也从早期的平台中介撮合,逐步过渡到数据产品、数据服务等全方位配套的交易及服务模式。随着整个数据交易行业的发展及规模的增长,交易将形成更

深的行业渗透和更广的行业应用范围。可以预见到数据的交易模式的演进,将会出现交易模式的细分,同时也会涌现出一些混合模式和混合业态。同样,就像传统的商品交易市场的演进一样,也会出现交易代理、交易中介机构,同时,基于数据产品和数据服务的衍生市场,如期货、期权等二级乃至三级市场都会逐步发展出来,会有一个欣欣向荣的数据交易生态的形成。而与其紧密相关的政策支持、法律法规、行业监管、数据开放和共享、交易模式和方法的演进、数据价值的深度发掘都是不断进化的主题和挑战。

#### 4. 数据资产安全

数据资产管理的目的是保障数据资产安全,数据交易的前提也是数据资产的安全。当前数据资产安全面临多方面的挑战,主要原因为法律法规不健全、市场缺乏信任机制、市场参与方众多、缺乏完善的数据资产管理手段及技术等。一方面,对于数据的隐私及安全的法律法规都还在制定和实施过程中;另一方面,数据又需要在市场有效地流通。数据所有者、数据交易中介、技术服务方等都可能私下复制并对外共享交易数据,数据使用者不遵守协议要求,私自留存、复制甚至转卖数据的现象也普遍存在,数据交易中心的平台及交易系统存在安全隐患,这些问题都有可能出现。

按照国家或机构的法律法规及行业监管要求,通过评估数据资产安全风险,制定管理制度规范保证数据资产安全,建立良性互动的数据交易生态体系势在必行。面对复杂的数据资产管理、交易及流通环境,亟须建立包括政府、监管机构、社会组织等多方参与的、法律法规和技术标准多要素协同的、覆盖数据资产交易流通全过程和数据全生命周期管理的数据资产管理及监管生态体系。

## 5.3 数据要素流通及市场化配置

### 5.3.1 数据要素化

在数字经济时代,数据的一个根本特性是它的生产要素性。数据之所以能起到革命性和颠覆性的作用,最根本原因就是数据成为一种新型生产要素。数据由于本身所附带或隐含的价值,被类比为新时代的石油、黄金,而且被正式认同为“一种与资本与劳动力并列的新的生产要素”。也就是说,数据不仅对生产过程中形成产品和产生价值起着重要的作用,其本身更是作为像资本和劳动力这样的生产要素,是产品生产不可或缺的元素,也是最终产品中不可分割的一部分。

生产要素有劳动力、资本以及土地等自然资源。传统的生产方式是通过劳动力加工自然资源,把它们变成产品进行销售,在其中产生增值。当数据成为一种生产要素加入生产过程时,可以完全替代其他原有生产要素,或是改变原有要素的构成比例。一个简单的例子就是自动驾驶,通过学习和掌握人类的驾驶行为,使用传感器和基于人工智能的自动驾驶软件,可以完全替代最有经验的司机,在这里不再需要司机这一劳动力要素了,这样就整个颠覆了出租车行业和驾驶行业。再如互联网金融,在缺乏数据的情况下,一个传统的银行要放贷的话,需要对贷款的企业进行线下调查,如经营状况、员工数量、固定资产、有没有资产抵押等,再进行各种各样的分析,可能需要一个多月才能放一笔贷款,即使这样也不能保障这个企业可以顺利还款。而通过电商平台上面的数据,可以了解商户所有的业务、资金周转、信用等情况,放贷只需要几分钟甚至更短的时间,大幅地节约放贷成本和周期,这就是数据成

为生产要素,不需要那么多的时间、那么多人力以及资金成本来决定是否放贷。传统银行很难与这种新兴的基于数据做征信和风控的新型互联网银行竞争,面临被淘汰出局的危险。数据作为新的生产要素,正在变革全行业的格局。

当前,数字经济的发展核心就是数据价值的发挥。数据作为数字经济建设的关键要素,将对其他生产要素产生替代或是倍增效用,为经济转型发展提供新动力。数据要素化该如何实现?如前所述,第一是数据的资源化,涉及原始数据的获取以及数据后期的加工组织,构建数据价值释放的潜力。当前,数据作为基础性、战略性资源已经得到广泛共识。第二是数据的资产化,数据变成可直接变现的产品、商品和服务,使得数据价值可以度量、可以交换和交易。第三在资产化的基础上实现资本化,数据资本化就是把数据作为一种金融资产,可以用于借贷、抵押、融资等,数据资本化使数据由货币性资产向可增值的金融性资产转换。以此让数据要素价值得以全面发挥和释放,并创造巨大的新价值。

### 5.3.2 数据要素流通

如上所述,数据生产要素化之后,需要流通和交易才能实现其价值,并在此过程中可能创造出新的价值。数据要素流通的全过程非常复杂,包含有多达十几个不同的阶段,可以大体上分为3个阶段。

第一阶段包含数据确权、数据加密、数据存储、数据水印。这个阶段首先是确定和保障数据的原始权属及数据安全,所以在确权的基础上,还需对数据进行加密,并做好数据的存储及备份管理。另外,为了在数据的使用过程中能够进行追踪、溯源和防止泄露及滥用,还需要一定的数据水印及溯源跟踪技术。

第二阶段包含数据授权、数据评估、数据定价、数据分析、数据

融合、数据交换、数据交易,这一阶段构成了数据要素流转、交易及数据要素市场的基础,数据在有效的授权及托管体系下,可以经过分析、加工及融合,再经过交换,流通到数据的需求方及消费方手中,在此过程中还可以进行评估、定价、撮合、交易,实现有效的价值分配。数据的分析、加工、融合、交换的过程能够以多种方式进行组合,这个过程还能形成迭代。如医疗数据,经过一些数据源的初步融合之后,需要多个分析、加工过程的迭代,其间再融合进其他数据源及加工过的数据,最后才能形成产生很大商业价值的医疗数据产品,其中的交换及交易过程也会非常的复杂。

第三阶段包含数据的应用及数据血缘。这一阶段是数据的最终应用及价值体现。数据可以应用在各个行业及产业体系中,与应用结合产生价值。数据血缘是一个专有名词,可以从数据的最终应用及价值体现中回溯到原始数据及所有加工的过程,分析哪些原始数据产生了最终的应用价值,也能够依据整个过程对数据的价值构成及价值分配提供有效的参考,促成合理的评估定价体系及价值分配体系。

数据要素的流转在数字经济时代是一个创新前沿的领域。尽管大家都在探索,但由于它所涉及的参与方及流程的复杂性,目前还没有完整系统的解决方案及最佳实践指南。在数据应用场景不断发展、不断扩大的未来,可以逐步将这个领域的流程、规范、标准及相关体系建设并完善起来。

### 5.3.3 数据要素市场化配置

数据要素的市场化,首先是对数据要素确定其原始的权属,并对数据要素使用的全过程进行确权、存证、跟踪、溯源、管理,在保障数据的安全及隐私的基础上,促进数据要素的流转,同时还能建立多方参与的价值分配体系,整体上明确数据的责、权、利,实现有

效的数据要素市场配置,最大限度地发掘数字要素的市场价值。

随着数据资产化发展,数据要素的价值得到不断的开发和拓展,数据要素市场也逐步形成了数据交易主体、数据交易媒介、数据交易监管的市场格局。前面也强调过,数据资产确权是数据要素市场培育和发展的重点和难点。数据权属问题是数据流通、交易的核心问题,数据权属不确定,导致数据流通、交易、使用不能被规范化,相关主体的权责不清,数据市场价值开发处于灰色领域,数据监管困难,成本增加。因违法成本较低,个人信息在黑市上低价售卖、隐私泄露问题等数据违法行为频发。

数据产权制度的建设和完善,对于数据要素权益分配制度的建设具有重要的意义,如确立数据资产的所有权,事实上是明确了数据要素收益的所有权归属问题。除此之外,数据的管理权、转让权、使用权、知情权等事实上都明确了数据在不同经济活动、不同环节的权益合法性和归属问题。

需要指出的是,若要明晰数据权益的所属关系,关键在于做好数据权利分割、数据分类和数据分级,并根据数据的类型、数据的特性,分级、有区别地进行精准化管理,对于重要的、安全要求高的国家数据或者企业数据,可以不公开不共享。对于较重要的、安全要求较高的数据,可以有条件地共享和开放,采用隐私计算或区块链技术,实现数据“可用不可见”“可算不可识”,而对于那些具有公用特性的数据可以采用数据集或者 API(应用程序接口)的形式开放共享。

在明确好数据要素的责、权、利之后,数据要素的市场化配置,重点在于建立公平、高效的数据流通及收益分配体系,促使数据要素朝着最能产生市场价值、最符合市场需求的产业及方向流动,避免出现数据浪费、数据无效流转、数据垄断、数据效益分配不均等不良现象。要建立体现效率、促进公平的数据要素收益分配体系,一方面需要加强数字信息基础设施建设,开放竞争有序的数据要



素市场,加强数据交易和管理平台建设,促进数据要素在地区、行业、企业、部门之间的流通;另一方面,需要培养多元数据要素市场主体,创新数据要素的商业模式、交易模式、评估定价模式、收益分配模式,推动数据要素的价值开发。

此外,还需创新数据要素流通技术和手段,加强数据技术人才队伍建设,促进隐私计算、区块链、差分隐私、数据标识等数据安全技术在数据交易和流通中的应用,推广“数据可算不可识”“数据可用不可见”模式,利用技术手段保障数据交易和流通的安全性。

通过所有这些体系、手段和方式的结合,最终能够看到数据要素发挥其最大的生产力价值,促进全球和全社会的数字经济建设及发展。