

第 1 章

数据与数据资产

1.1 数据与大数据

“数据”是个外来词，来自拉丁语“datum”^①，意思是“给定或授予的事实”。古典用法中，数据指与数学计算或统计相关的数字；但在现代社会中，数据的内涵远超数字。专家估计“数据”这个词大约出现在 17 世纪，但远早于 17 世纪，人们就意识到数据的重要性。远古时代，在文字和数字出现之前，先民们就已经开始用在绳子上打结的方式来记录数量，这就是著名的“结绳计数”，这个数量也是最早的数据。人类社会生活中非常重要的数据是个人、家庭、团体的财产收入和支出记录，从国家层面来看，这就是财税记录。无论是古代王朝还是现代国家，都需要财税记录的详细数据，以便做出适当的财政安排。由于土地是重要的税收基础，因此土地数据是财税数据的核心要素。英国最早的财税资料是 1087 年威廉一世清查全国土地所得的结果，被记录在《末日审判书》

^① 数据的英文是 data，是 datum 的复数形式。

中。此外，英国财政署的《财政署收入年表》收录了自 12 世纪至 19 世纪每年英格兰各郡郡守交纳给国王的赋税和拖欠金额的资料，现存最早的资料是亨利一世 1130 财年的记录。

除土地数据外，人口数据也非常重要。按人口征收的税称为“丁税”或“人头税”。国家要按人口征税，就需要掌握现有人口数量和变化的数据。除了丁税，还存在兵役、徭役等与人口相关的非货币性税负，这就需要人口的年龄、性别等数据，因此需要进行更详尽的人口普查。最早与现代人口普查接近的数据是中国明朝初期（14 世纪）官府登记户口的簿册，每户的分册称作“户帖”，由居民保管；辖区内各户的总册称作“户籍”，留归官府并报户部。户帖调查项目包括五项：户的种类（分军户、民户、匠户），户主籍贯（原籍、现籍），居住地（乡、都、保、圩），家庭成员（姓名、年龄、人数），与户主的关系。图 1.1 是明朝洪武四年（公元 1371 年）官府发给徽州府祁门县汪寄佛家的户帖。值得注意的是，户帖不仅包含了数字信息（比如家庭成员的年龄、人数），也包含了大量文字信息（比如户的种类、户主的籍贯等），这说明数据的范围远超数字。



图 1.1 中国明朝初期的户帖

无论是个体、企业还是国家，拥有全面、完整、高质量数据的重要性都是显而易见的。但在古代，获取高质量的数据是很困难的，为应对数据匮乏而采取的一些措施常会带来意想不到的结果。欧洲历史上征收过一种窗户税，其本质是房产税，但因为当时的政府无法获得关于房产价值的高质量数据，只能用易获得但低质量的数据——窗户的数量来代替。窗户越多的房产被认为价值越高，所需缴的税也越多。英国于1697年开征窗户税后，人们想方设法地减少房屋的窗户数量以减少纳税金额，甚至逃避纳税义务。为了多征税，英国不断下调起征的窗户数目，而英国房屋的窗户数量也随之减少。比如1776年，当起征窗户数目从10扇下调到7扇时，英国几乎2/3的房屋的窗户数量都减少到6扇。虽然减少窗户数量可以少纳税，但屋里的阳光和清新空气也同时减少，人们被迫生活在阴暗、潮湿、空气浑浊的房子里，不仅损害了健康，也导致传染病更易流行，社会的公共卫生状况急剧恶化。在社会的巨大压力下，英国于1851年废除了窗户税。法国在大革命期间从英国引进窗户税，但直到第一次世界大战快结束的1917年才废除。窗户税被废除了，但房产税并没有废除，实际上房产税是现代欧美国家的重要税种。由于政府掌握了更准确的房产价值的数据（建筑年代、建筑标准、建筑面积、交易记录等），政府可以更高效地征收房产税，同时又不会造成窗户税那样的公共卫生悲剧。窗户税的历史充分说明了获取高质量数据的重要性。

早期的数据主要以数字和文字的形式存在，但随着信息技术的发展，数据逐步演化成对可传输和可储存信息的总称。现代社会中数据的概念不仅包含数字和文字，也包括图像、声音、视频等形态的信息。例如，2018年4月7日至6月9日，张学友在南昌、赣州、嘉兴、金华举办了4场巡回演唱会，警方在这4场演唱会上总共抓获了5名逃犯。警方并不是事先知道逃犯要去看张学友的演唱会，而是通过演唱会安检时的人

脸识别与逃犯数据库进行自动比对时发现的。同样的技术也被广泛应用在火车站、机场、汽车站、商场、酒店等人流密集的地方，让逃犯无处藏匿。在这个例子里，对抓捕逃犯起到关键作用的是海量的人脸图片数据以及高效的人脸识别技术。

在万物互联的时代，所有能够被记录和存储的信息都将被数字化（digitized），从而成为数据。随着信息技术的高速发展，现在人类每天产生的数据是一个天文数字，并且这个天文数字还在快速上升。以照片为例，人类的第一张照片拍摄于 1826 年，到 185 年后的 2011 年人类拍摄的所有照片的总数估计为 3 800 亿张。华为最新发布的《华为影像 XMAGE 2023 年度趋势报告》显示，2023 年全球拍摄的照片总数可能超过 1.4 万亿张，其中 89% 以上通过手机拍摄。换句话说，2023 年一年新拍摄的照片数量是自照片问世后的 185 年（1826—2011 年）拍摄的所有照片数量的 3.68 倍。

根据美国著名的全球数据和商业信息平台 Statista 的估计，人类在 2023 年产生的数据总量大约是 120 ZB（一部时长 2 小时的高清电影的数据量通常不超过 0.01TB，而一个 ZB 是 10 亿个 TB）。超过一半的互联网数据由视频贡献。在 2010 年时人类产生的数据总量才约 2 ZB，13 年后这个数字翻了 60 倍。该机构估计 2025 年产生的数据总量可能会达到 180 ZB^①。全球历年的数据总量如图 1.2 所示。

2012 年 12 月牛津大学教授维克托·迈尔·舍恩伯格（Viktor Mayer-Schönberger）与《经济学人》杂志的数据编辑肯尼斯·库克耶（Kenneth Cukier）合作出版了《大数据时代》（*Big Data: A Revolution That Will Transform How We Live, Work, and Think*），从此，“大数据”这个新名词开始被广泛使用和传播，逐渐成为日常用语的一部分。顾名思义，大数

① <https://www.statista.com/statistics/871513/worldwide-data-created/>.

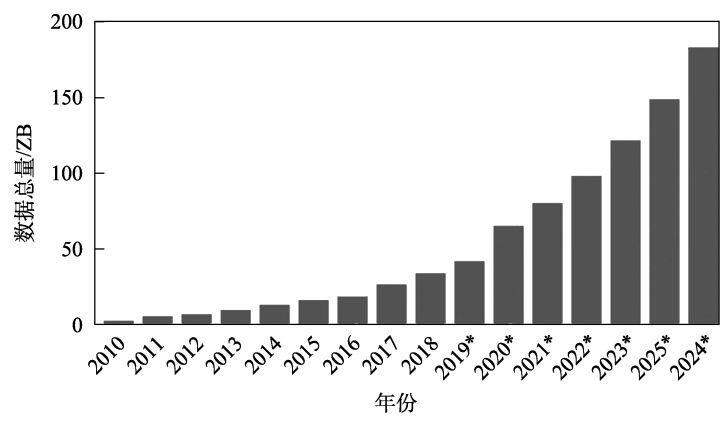


图 1.2 全球历年的数据总量
(注：图中 2019 年至 2024 年数据为 Statista 预估。)

据指规模庞大的数据。根据维基百科的定义，大数据是一种大规模的数据集合，其规模远远超过传统设备、数据库和软件的存储及分析能力。我们认为大数据有四个重要特征：体量大、速度快、种类多、价值密度低。在商业领域，比如市场营销、风险管理、人力资源管理等方面，大数据可以帮助企业更好地理解客户行为，从而优化产品和服务。在社会领域，比如公共安全、医疗保健、城市规划等方面，大数据可以帮助政府改善决策，更好地提供社会服务。在教育领域，大数据可以帮助学校和老师研究并分析学生学习参与、学习表现和学习过程的相关数据，进而对课程、教学进行实时修正并预测学习者未来的学习趋势，有助于实现弹性学习、个性化辅导，真正做到因材施教。在医疗领域，大数据应用为医疗科学研究和临床实践带来了巨大的便利，帮助医生更准确地诊断和治疗疾病，帮助医疗企业更高效地研发新的药物和生产新的器械，提高医疗质量和效率，并降低医疗成本。大数据是一场革命，深刻地改变了我们生活、工作和思考的方式。在《大数据时代》出版 11 年后的今天，大数据的价值已经被广泛认可，并以前所未有的速度影响和塑造着

我们的生活。

在简单回顾了数据和大数据的历史后,接下来我们将从财务的角度,深入讨论什么样的数据可以被看作是一种资产,并计入财务报表。我们强调,并非任何数据都有显著的经济价值,数据也不是越多越好。对数据资产的认定、评估和入表需要有深厚的理论基础、严格的判定标准和完备的操作流程。

1.2 数据资产的定义与特性

尽管数据在经济和社会活动中的重要性早已被广泛认可,但数据资产却是一个新概念。数据资产这一新概念产生的背景是,以大数据、云计算、区块链、物联网、工业互联网、5G、人工智能等为代表的新一代信息技术加速突破应用,从而带来数字经济的蓬勃发展。根据中国信息通信研究院发表的《全球数字经济白皮书(2023年)》的估算,2022年全球51个主要国家或地区的数字经济规模为41.4万亿美元,同比名义增长7.4%,占GDP的46.1%。从规模看,美国数字经济规模达到17.2万亿美元,蝉联世界第一,占GDP的比例达到67.5%。中国的数字经济规模为7.5万亿美元,位居世界第二,占GDP的比例约为41.5%。毫无疑问,数字经济已成为这个时代最强大的增长引擎。数字经济的背后是庞大的数字基础设施建设,以及人工智能在各个领域的深入应用。数字经济的核心基础是大数据,没有大数据,再好的算法也无法产生人工智能,也就无法提升生产力,无法创造价值。

在2019年之前,公认的生产要素包括土地、劳动、资本和技术。2019年10月,我国首次将数据纳入生产要素。这是非常重大的一项举措,反

映的正是数据融入了生产、分配、流通、消费等各个环节，深刻地改变了经济发展驱动力和社会运行方式的现实。数据和传统生产要素结合可以优化资源配置、提高生产和管理效率、促进生产过程的自动化和智能化。数据的应用可以帮助企业发现新的商业机会和价值点，创新业务模式和产品服务，提高企业的竞争力和盈利能力。随着大数据的不断累积和完善，以及人工智能等技术的不断进步，企业可以更加深入地挖掘数据的价值，实现数据驱动的决策和生产。数据的价值在于其能够为企业提供深入的洞察力，优化决策过程，提高运营效率，甚至创造新的商业模式和市场机会。这种价值创造是数据成为资产的基础前提。

在讨论数据资产之前，我们首先要了解什么是资产，然后再分析数据资产区别于其他资产的特性。资产是指企业过去的交易或事项形成的、由企业拥有或控制的、预期会给企业带来经济利益的资源。从资产的定义中我们可以看出资产有两个基本属性：第一，资产必须是企业拥有或控制的资源，即资源的经济所有权是明确的。企业可以使用但没有所有权或控制权的资源不是该企业的资产。例如，飞机和机场对航空公司都非常重要，但航空公司只拥有或控制飞机，对机场只是付费使用，没有所有权或控制权，因此飞机是航空公司的资产，而机场不是。第二，资产的核心特征是预期会给企业带来经济利益，不能带来经济收益的资源则不能视作资产。例如，对于一家运输公司来说，能够正常行驶的车辆是资产，但完全报废的车辆则不是资产。

数据作为一种与传统资产不同的信息资产，如果要纳入资产核算范畴，需要同时满足资产的两个基本属性：具有明确的经济所有权和收益性。学术界对数据资产需要满足资产的这两个基本属性没有争议。比如，司雨鑫（2019）将数据资产定义为企业日常运营中形成的、由企业拥有或控制的、以数据形式存在的、预期可以带来现金流流入的非货币性资

产。张俊瑞等（2020）将数据资产解释为由企业拥有或控制的、具有数据化形态的、可辨认的非货币性资产。许宪春等（2022）认为数据资产是拥有应用场景且在生产过程中被反复或连续使用一年以上的数据。按照《中国国民经济核算体系（2016）》和《企业会计准则——基本准则》对于资产的定义，同时参考其他学者对数据资产的定义，我们将数据资产定义为“企业拥有或控制，预期会给企业带来经济利益的、以数据为主要内容和服务的可辨认的非货币性资产”。

数据作为生产活动的结果，从产品类别看，它兼具货物和服务的特征^①。由于数据本身的种类和性质较为复杂，如其应用场景的价值多元性、市场价值的不确定性、二次开发的价值叠加性等，使得数据资产也具备场景依附性、非消耗性、时效性、共享性和非竞争性等显著异于传统资产的特性^②。

场景依附性：数据资产的价值与应用场景相关，同样的数据资产被不同企业应用会创造不同的价值。不同行业的企业由于其自身的经营场景和管理模式不同，对数据的需求和定价也不同。例如交通出行数据，网约车公司可以利用它来匹配司机和乘客，提高出行服务效率，增加收入；政府部门则可以利用这些数据优化道路建设，提高交通服务质量。显然，同样的数据对网约车公司和政府部门的价值是不一样的。但传统资产则没有这个特性。比如车辆，无论它是属于网约车公司还是属于政府部门，其价值都主要由车辆本身决定（取决于型号、里程、年限等），而与应用场景关系不大。因此，数据资产的一个重要特性是其价值依附于应用场景，而应用场景又具有多样性，这就使得对不同应用场景下的数据资产进行归类、管理并统一评估成为一项重要问题。

① Mandel, 2012。

② 许宪春等，2022。

非消耗性：数据资产的价值虽然可能会随着企业的更新加工、不断积累挖掘或特定事件的发生而产生变化，但不会因为被使用而产生减值或损耗^①。反之，传统的实物资产是具有消耗性的，传统资产的使用会带来折旧。数据资产非消耗性的特征给数据资产的折旧核算带来挑战。

时效性：新数据产生、社会经济环境变化或企业商业模式变更等内外部因素，都会导致原有数据资产升值或贬值，价值变化的速度主要取决于数据的类型和用途。比如，金融市场上的价格数据是数据时效性最显著的代表。因此，数据资产的价值与其评估时间点有着紧密的联系，其市场价值的波动对于价值评估而言是一大难点。

共享性和非竞争性：数据资产本质是一种数字信息，独立于存储的物质载体，即它能同时在不同的物质载体上存储（比如同时存在于云端和不联网的服务器中），可以实现多主体同时共享，与无形资产的共享性有相似之处。但区别于传统的无形资产，数据具有非竞争性，一种数据被一个人获得后，其他人还可以继续获得该数据，数据复制的边际成本几乎为零。因此，数据资产的共享并不会给使用主体带来损失，其价值也不会因重复销售而减损。而传统无形资产则明显具有竞争性，比如专利，专利存在的目的正是要确保竞争对手无法使用这项资产。

数据资产的形式多种多样，其分类与数据资产的会计计量和估值息息相关。比如，可以按照价值属性、生产对象、存储形式、权属、供给方式以及使用方式等对数据资产进行分类：

（1）数据资产按价值属性可分为资源性数据资产和经营性数据资产。资源性数据资产是数据资源在尚未进入流通市场之前的形态，主要是在内部形成并使用；经营性数据资产则是指被产品化以后的、可以直接在市场上作为“商品”流通的数据资源。这一分类方法也是数据资产

^① Rassier 等，2019。

化实施路径的重要依据。

(2) 数据资产按生产对象可分为与人、货或者场景有关的数据。与人相关的数据指围绕个人生产生活所产生的数据，比如个人的年龄、性别、收入等；与货有关的数据包括与产品、设备等实物相关的信息，如产品的型号、性能等；与场景有关的数据指人与周围环境的连接数据，如家庭生活起居信息、公交车日客流量等。

(3) 数据资产按存储形式可分为结构化数据、非结构化数据和半结构化数据。结构化数据指严格遵循数据格式和长度规范，主要通过关系型数据库进行存储、调用和管理的数据；非结构化数据可以理解为不满足上述条件的数据，比较典型的就文本、图像、视频等；半结构化数据则居于两者之间。

(4) 数据资产按权属可分为公共数据、企业数据和个人数据。公共数据指由政府及相关单位在提供公共服务的过程中产生的数据；企业数据则是企业在生产经营活动中产生的数据；个人数据顾名思义，即由个人在日常生活中产生。

(5) 数据资产按供给方式可分为自给性数据与交易性数据。自给性数据产生于内部，而交易性数据来自外部。

(6) 数据资产按使用方式可分为供国内市场生产者使用、供国内非市场生产者使用，以及供出口国外使用，其中，国内非市场生产者主要包括政府与非营利性组织。

1.3 数据资产和财务报表

由于数据资产具有场景依附性、非消耗性、时效性、共享性和非竞

争性等特性,因此数据资产在价值分配与计量上与传统资产有较大差异,对企业数据资产的估值也产生了一定挑战。尽管如此,确认、评估和计量数据资产在技术上依然存在很多困难和争议,但学者们普遍支持将数据资产化。《大数据时代》中写道:“虽然数据还没有被列入企业的资产负债表,但这只是一个时间问题。”在该书出版的2012年,数据资产入表还只是一个设想,但在2024年,这已经成为现实。

在2019年10月数据首次被纳入基本生产要素之后,我国出台了一系列与数字经济,特别是大数据相关的政策法规,全方位覆盖了宏观层面数字经济的顶层设计、中观层面数据要素的制度安排以及微观层面数据资源的规范建立。2021年3月,《中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标纲要》做出了完善数据要素产权性质、建立数据资源产权相关基础制度和标准规范、培育数据交易平台和市场主体等明确的战略部署。同年6月,十三届全国人大常委会通过了《中华人民共和国数据安全法》,这是我国首部数据安全领域的基础性律法。2022年12月,《中共中央国务院关于构建数据基础制度 更好发挥数据要素作用的意见》正式出台,标志着我国数据要素市场进入了有序规范的探索和发展阶段。

在对数据资源的会计处理方面,我国也提出了相应的制度安排。2023年8月23日,财政部发布了《企业数据资源相关会计处理暂行规定》(以下简称《暂行规定》),自2024年1月1日起开始实施。《暂行规定》根据《中华人民共和国会计法》和企业会计准则等相关规定,首次明确了数据资产的确认、适用范围、会计处理标准以及披露要求等内容。具体来讲,《暂行规定》按照数据资源有关的经济利益的预期消耗方式,根据企业持有对客户提供服务、日常持有以备出售等不同业务模式,将数据资源分类为无形资产和存货科目进行确认、计量和报告。这一规定的颁

布对规范企业数据资源相关会计处理和加强相关会计信息披露具有重要意义。《暂行规定》首次从政策角度明确允许将数据资产确认入表，使得原先只能费用化处理的数据资源开发成本在满足一定条件后得以确认为资产，为报表使用者提供决策有用信息，同时帮助数据驱动型企业吸引外部融资、优化财务结构、提升公司价值。我国关于企业数据资源的相关探索不仅有助于监管部门完善数字经济治理体系，还有助于我国在国际会计准则制定等工作中贡献中国智慧、提供中国方案。为了解决数据入表实践当中的一些困难，进一步优化数据资产入表的工作，财政部于2024年1月11日印发了《关于加强数据资产管理的指导意见》（以下简称《指导意见》）。《指导意见》在数据资产管理、数据资产权责确认、数据资产开发使用、数据资产价值评估、数据资产披露等方面都做了清晰阐述。

数据资产入表对于数据要素型企业（也称“数商”）的影响最为显著。根据上海市数商协会的分类，数商主要有以下四类：

（1）资源供给型数商：以提供数据资源作为主要业务形态，通过收集自身业务活动产生数据或集成外部数据等方式持有数据资源，形成数据产品并对外提供；

（2）技术赋能型数商：以提供技术工具或技术服务为主要业务形态，具备数据存储计算、加工处理、分析挖掘、治理、安全等领域的成熟技术能力，并对外提供相应的技术类支持与服务；

（3）生态服务型数商：以提供除数据资源、技术能力外的其他服务为主要业务形态，能够提供数据合规、安全、质量等评估服务，或者交易撮合、交易代理、专业咨询、数据经纪、数据交付等中介服务；

（4）数据消费型数商：以整合、采购数据资源并进行应用为主要业务形态，具备根据业务需求形成相应数据消费品，并在企业实际经营过

程中实施应用的能力。

资本市场对数据资产入表政策的反响是非常积极的。在《暂行规定》政策发布当天，数据要素板块大涨。截至当日收盘，国家发改委数据合作平台上海钢联、数据分析与决策支持服务机构零点有数、大宗商品价格指数提供商卓创资讯等多只个股涨停，人工智能及大数据服务商汇纳科技、数据智能服务商每日互动、数据资产化服务商易华录等涨幅超15%。随后一周，A股数据要素板块持续领涨。从资本市场对数据资产入表与披露政策出台的热烈反响中我们可以看出，投资者预期政策落地后相关企业数据要素价值将进一步释放，数字经济发展将进一步加快。