

第1章 数据与数据治理

数据（data）作为重要的生产要素已经融入人类社会生活的方方面面。习近平总书记多次强调，要构建以数据为关键要素的数字经济，在创新、协调、绿色、开放、共享的新发展理念指引下，推进数字产业化、产业数字化，引导数字经济和实体经济深度融合。可见，数据要素对经济社会的发展起着关键作用。伴随着科技与社会的发展，数据在被应用到各行各业的过程中也出现了一些问题，为了不断激发数据的核心价值，促进服务创新，数据治理（data governance）成为必由之路。本章的重点内容是对数据的价值、数据治理的内涵以及数据治理的发展进行介绍，为数据治理框架体系等内容学习进行基础铺垫，对数据治理相关内容提供整体把握思路。

【教学目标】

介绍课程的主要特色，使学生对本课程的整体设计、教学目的、核心内容等进行充分的理解。讲授数据与数据治理的相关概念、内涵、发展，通过课堂互动交流进行问题辨析，引导学生以多元化视角观察数据，注重数据作为发展要素的意义，对数据治理的基础进行深刻的认识。

【课程导入】

（1）NASA（National Aeronautics and Space Administration，美国国家航空航天局）如何能提前预知各种天文奇观？创业者如何确定自己的产品定价与服务对象？在未来的城镇化建设过程中如何打造智能城市？这一系列问题的背后，其实都隐藏着数据的身影——不仅彰显着数据的巨大价值，更直观地体现出数据在各个行业的广阔应用。那么数据在我们的日常生活中都以怎样的形式出现？

（2）数据在生活中无处不在。过去，人们习惯把数字的组合称为数据，但在今天，这样的理解显然不够全面。是否可以把数字、字母、符号等的集合称为数据？似乎信息与知识也是由这些元素组成的，那么数据、信息、知识之间存在怎样的关系？

（3）“啤酒与尿布”的故事源于在20世纪90年代的美国超市中，这两件看似毫无关联的商品会经常出现在同一个购物篮里。你认为“啤酒与尿布”蕴含了怎样的数据关联关系？

【教学重点及难点】

重点：本章的学习重点在于数据治理的内涵与数据治理的发展。首先要理解数据治理的概念，进而熟悉数据治理的相关内涵，对数据治理有整体的把握，为数据治理体系的学习打好基础。

难点：本章的学习难点在于对相似概念的辨析，如数据、信息与知识，数据治理与数



据管理。数据治理的发展部分涉及政策内容，时空跨度较大，与时势联系紧密，也是需要重点理解掌握的内容。

1.1 数据的价值

当前信息化、大数据、数字经济等高频词语已成为世界各国推动经济社会可持续发展的着力点和竞争点。《经济学人》杂志曾将数据比喻为“21 世纪的石油”，数据的重要性不言而喻。本节主要介绍数据的价值，包含数据的相关概念以及数据创造价值的维度。数据是通过观测得到的数字性的特征或信息。随着社会的不断发展进步，数据已经成为重要的生产要素，世界各国都将其作为国家战略不可或缺的组成部分，数据的价值不容忽视。本节通过介绍数据对于个人、企业与国家 3 个层面的价值以帮助读者更好地理解数据的重要意义。

1.1.1 数据相关概念辨析

1. 数据的概念

根据最新版的《牛津英语词典》，数据是“被用于形成决策或者发现新知的事实或信息”。根据国际标准化组织（International Organization for Standardization, ISO）的定义，数据是对事实、概念或指令的一种特殊表达方式，用数据形式表现的信息能够更好地被用于交流、解释或处理。在《现代汉语词典》（第 7 版）中，对于数据的解释是：“进行各种统计、计算、科学研究或技术设计等所依据的数值。”数据是人类通过观察自然、科学计算或社会实践等多种方式得出的一种记录，是对人类社会的一种描述、记录和表达。数据有许多形式的载体，最简单的一类就是数字，也可以是符号、文字、图像、声音、视频等。通常，在学术研究论著中，数据只是作为“信息的单元”。在数字经济时代，数据被纳入新型生产要素，是人类社会发展进步的基础。数据成为反映环境、优化体系、驱动发展的重要元素。在科学研究、商业管理（比如涉及销售、收入、利润、股价等）、金融、政治（比如涉及犯罪率、失业率、识字率等）和事实上的其他一切人类组织性活动形式（比如非营利性组织所做的流浪人口调查等）中，数据无处不在。

2. 信息的概念

信息是物质存在的一种方式、形态或状态，也是事物的一种普遍属性，一般指数据、消息中所包含的意义，可以使消息所描述事件的不定性减少。美国数学家、信息论的奠基人克劳德·艾尔伍德·香农（Claude Elwood Shannon）在他的著名论文《通信的数学理论》（1948 年）中提出计算信息量的公式为

$$H = -\sum_{k=1}^n p_k \log_2 p_k \quad (1.1)$$

式中， n 为组成信息的符号数； p_k 为符号 k 出现的概率。

从式（1.1）可知，信息量计算公式的意义和热力学中熵的本质一样，故信息量也称为信息熵。同时，当各个符号出现的概率相等，即“不确定性”最高时，信息熵可以被视为对“不确定性”或“选择的自由度”的度量。美国数学家、控制论的奠基人诺伯特·维纳（Norbert Wiener）认为，信息是“我们在适应外部世界，控制外部世界的过程中同外

部世界交换的内容的名称”。英国学者阿希贝（W. R. Ashby）认为，信息的本性在于事物本身具有变异度。意大利学者朗高（G. Longo）在《信息论：新的趋势与未决问题》中认为，信息反映了事物的形成过程、关系及差别，包含于事物的差异之中，而不在事物本身。

信息与数据既有联系，又有区别。信息是数据的内涵，是加载于数据之上，对数据作出的具有含义的解释。数据和信息是不可分离的：数据是符号，是物理性的，信息是对数据进行加工处理之后所得到的并对决策产生影响的数据，是逻辑性和观念性的；数据是信息的表现形式，信息是数据有意义的表示。

3. 知识的概念

知识是对某个主题确信的认知，意指透过经验或联想，能够熟悉进而了解某件事情，这种事实或状态就称为知识，其包括认识或了解某种科学、艺术或技巧。知识不是数据和信息的简单积累，知识是可用于指导实践的信息，知识是在人们改造世界的实践中所获得的认知和经验的总和。知识分为显性知识和隐性知识。显性知识是已经或可以文本化的知识，并易于传播。隐性知识是存在于个人头脑中的经验或知识，需要进行分析、总结和展现，才能转化成显性知识。知识也是人类在实践中认识客观世界（包括人类自身）的成果，包括事实、信息的描述或在教育和实践中获得的技能。知识是人类从各个途径中获得的经过提升总结与凝练的系统性认知。

数据、信息和知识都是社会生产活动中的基础性资源，都可以采用数字、文字、符号、图形、声音、影像等多媒体来表示。三者都具有客观性、真实性、正确性、价值性、共享性等特点。数据、信息和知识都是对事实的描述，被统一到对事实的认知过程中。首先，由于人类认知能力存在一定局限性，或者人类认知世界过程中所采用的工具受到约束，导致了数据只是对事实的初步认知甚至存在错误认知的表述；其次，人类借助思维的发散性以及信息技术等对数据进行处理，进一步揭示事实中事物的联系，从而形成信息；最终，在实践中，经过周而复始的处理与验证，现实中事物之间的关系被正确地表示出来，因此形成了知识。

通过数据、信息及知识的概念可以看出，孤立的数据是没有意义的，当数据被置于情境之下审视或经过分析处理之后，“数据”就会变为“信息”；一般而言，数据经由处理后称为信息，从这些信息中分析出来的讯息称为知识，再通过不断地行动与验证，逐渐形成智慧。举例来说（如图 1.1 所示），“10”就是一个客观的记录，这个数据本身无意义；但是“10 个零件”就是有意义的信息，此时的数据可以称为信息；数据的价值还可以进一步体现为知识，例如，“生产线每天生产能力为 10 个零件”，此时的数据融合了固化的经验，可以称为知识；数据的价值最高可以体现为智慧，例如，“通过流程优化将生产线每天生产能力提升到 20 个零件”，此时的数据是组织解决问题能力的基础，是最具价值的信息形式。

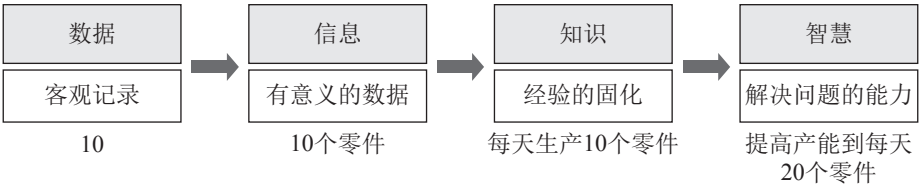


图 1.1 数据、信息、知识与智慧的关系



1.1.2 数据作为新型生产要素的主要特征

1. 数据是新型核心生产要素

人类社会已经从农业经济、工业经济进入到了信息经济时代，农业经济时代的核心生产要素是劳动力与土地，工业经济时代的核心生产要素是资本与技术，进入 21 世纪以来，以大数据、人工智能、区块链、量子计算等信息技术为标志的新一轮科技革命和产业变革悄然而至，数据量和算力呈爆炸性增长，数据成为驱动经济社会发展的关键生产要素。因此数字时代核心生产要素就是数据，如表 1.1 所示。

表 1.1 不同历史阶段生产要素变迁及代表人物 / 事件

历史阶段		生产要素	代表人物 / 事件
农业时代		土地、劳动	威廉·配第，欧根·冯·庞巴维克
工业时代	第一次工业革命	土地、劳动、资本	亚当·斯密，让·巴蒂斯特·萨伊，约翰·穆勒
	第二次工业革命	土地、劳动、资本、组织	阿尔弗雷德·马歇尔
数字时代		土地、劳动、资本、知识、技术、管理、数据	《中共中央关于坚持和完善中国特色社会主义制度推进国家治理体系和治理能力现代化若干重大问题的决定》

数据在数字经济发展过程中具有关键作用，对传统生产要素也产生了深刻影响，展现着其巨大的价值和潜能。作为生产要素的数据，其本身虽然不能被直接用于生产经济物品，但是能在生产过程中发挥作用，如创造新的知识或者形成未来的预测，进而指导经济物品的生产。2019 年 10 月，中国共产党第十九届中央委员会第四次全体会议审议通过的《中共中央关于坚持和完善中国特色社会主义制度推进国家治理体系和治理能力现代化若干重大问题的决定》中提出，健全劳动、资本、土地、知识、技术、管理、数据等生产要素由市场评价贡献、按贡献决定报酬的机制。这是国家层面首次提出数据可作为生产要素按贡献参与分配，数据作为新生产要素从投入阶段贯穿到产出和分配阶段，数据资源的重要地位得以确立。这反映了随着经济活动数字化转型不断深入，数据对提高生产效率的价值与作用凸显，成为最具时代特征的新生产要素，标志着我国正式进入了数据红利进一步释放的阶段，数据将作为生产要素参与到市场的投入、管理、产出、分配等各个阶段。随着科学技术的创新与发展，新型数字技术不断涌现，数据作为要素是一个新命题，有大量前沿问题需要研究。在数字经济全面推进的时代背景下，培育发展数据要素市场，更需要充分发挥数据作为关键生产要素的价值。

2. 数据生产要素与传统生产要素相比所具有的不同特性

数据要素呈现独特新特征。与土地、劳动力、管理、技术、资本等传统生产要素相比，数据无疑是生产要素大家庭的“新面孔”。数据生产要素超越了传统要素的基本属性、作用形态和增值方式，一跃成为数字经济时代占据领先地位的战略资源，究其原因，主要是与传统生产要素相比，数据生产要素具有以下几方面的不同特征。

(1) 数据生产要素具有虚拟性。数据生产要素是一种虚拟的、存在于数据库与互联网空间中的资源，目前大多数已有研究都将数据的虚拟性视为该生产要素的一项核心特征。

虚拟性是数据与其他传统生产要素（如劳动力、资本和土地）的最主要差异，也是知识、技术、管理和数据等新生产要素的主要特点。虚拟性的存在意味着数据必须以其他生产要素作为载体才能发挥作用。在当前的技术条件下，数据在大多数时候存在于信息与通信技术产品中，二者有效结合构成了全球经济增长的主要动力。对于数据等虚拟生产要素的依赖是数字经济的主要特点之一，也是数字经济与传统经济的主要区别。

（2）数据要素具有非稀缺性。稀缺性是指资源因无法满足人类社会不断扩大的需求以及无法实现空间上的均匀分布而呈现出的稀缺状态。然而对于知识、信息、数据等无形资产来说，因其累积迅速且不占用空间而具有非稀缺性。数据要素的这种非稀缺性并非单纯意味着数据无处不在、无时不在，而且具有两重特定含义：一方面，数据在共享过程中可以实现指数级增长，数据传播链条越长，数量便如滚雪球般急剧增长；另一方面，数据在条件允许的情况下可以重复使用、循环使用乃至无穷尽开发利用，因此其对推动经济增长具有倍增效应。

（3）数据要素具有强劲的流动性。流动性是生产要素产生价值的基本前提，不同生产要素的流动性程度具有天壤之别。土地流动性最弱，劳动力流动性次之，技术流动性适度，资本流动性较强，而数据流动性最为强劲。不同生产要素在流动过程中呈现形态迥异的特征，劳动力流动和土地流转既是市场经济“看不见的手”作用的必然结果，也是在限定条件下人们追求收益最大化的结果。技术流动的前提是其具有先进性，如此才能在不同国家、地域、行业内部（或之间）进行输出与输入。投机性和逐利性是资本流动的基本特征，考虑到资本在流动时常常与其他要素裹挟在一起发生裂变，其对经济社会的推动作用更加显著。作为 21 世纪的国家战略资源，数据要素呈现出流动速度更快、渗透程度更深、涉及领域更广的特点，成为生产要素大家庭中最具增长潜力和价值的“璀璨明星”。

（4）数据要素具有时效性。从数据的生成到消耗，时间窗口非常小。数据的变化速率包括处理过程越来越快。数据的重要程度与时间成反比，如超过规定时限，数据可能就变得一文不值。当企业注销后，再对该企业相关数据作出分析将毫无意义。如随着物联网的广泛应用，信息感知无处不在，信息海量，但价值密度较低，通过强大的机器算法可以更迅速地完成数据的价值“提纯”，在大数据时代，数据快速处理能力将助推企业抢占市场先机。

（5）数据要素具有非排他性。排他性是指某一事物在一定范围内排斥另一事物的性质。经济学视角下的排他性则指在技术上排斥他人使用的可能性，比如某人使用或消费一种物品时便阻止了其他人同时使用或消费该物品。由此可见，资本和劳动力要素明显具有排他性。数据本身具备的非常强大的复用效率使其可以按照既有模式在一定范围按照一定权限重复使用，因此数据生产要素彻底颠覆了传统生产要素的使用局限性，数据资源可以无限复制，供多个主体在同一时间不同地点使用，这些主体的使用不仅不会减损数据本身的价值，还能够从对数据的利用中发掘出更多的价值。

（6）数据要素具有产权模糊性。数据生产要素在产权归属上存在一定的模糊性，其所有权和产生的各项产出在企业与消费者之间的分配尚不清晰。消费者在使用互联网公司等企业提供的各项 ICT 产品和服务的过程中会产生大量数据。这些数据往往由企业直接收集和整理，消费者在客观上没有处置和使用这些数据的机会。数据产权并不像传统要素的产



权那样，其主体对其所有物享有包含所有权、使用权、收益权、处置权等权利。因此，采用传统要素的产权界定手段很难清晰识别数据在不同场景中作出的贡献并追踪其主体，根据其应用的方式、创造的价值配置处置、收益等相应的权利。



视频讲解

1.1.3 数据创造价值的维度

数据与人类息息相关，越来越多的问题可以通过数据解决。不仅在数据科学与技术层面，而且在商业模式、产业格局、生态价值等方面，数据的价值都能够带来新理念和新思维，涉及面包括政府宏观部门、不同的产业界与学术界，甚至个人消费者。数据正在以前所未有的程度引起社会、经济、学术、科研、国防、军事等领域的深刻变革。数据所创造的价值将更好地解决商业问题、科技问题，以及各种社会问题。

1. 数据对于个人的价值

对于个人而言，数据是一个人在学习、工作、生活中形成的经验、知识等，乃至在个人同意出让前提下的个人信息数据实质上也是个人的重要“资产”，是一个人生存与发展的保障和动力。数据对个人的价值归根结底是能帮助人们全面了解事物本质并形成正确的决策，具体表现在以下几方面。

1) 提高工作效率

数据引发的技术革命以提升效率为起点。通过数据的计算与分析，许多时间被节省下来，人们可以去从事其他如消费、创新等活动。数据的运用不断释放人类社会的巨大产能，医疗记录数字化、出行方式智慧化等多方面的进步都源于数据创造的价值，丰富了人类的现实生活，增加了人类认知盈余。

2) 提高洞悉能力

用数据可以完成对事物的精准刻画，帮助人们全面了解事物的本真面目。此时，数据发挥的作用在于减少信息不对称，帮助人们提高洞察力和获得新知识。在数据的支持下，人们实现了从“不知道”到“知道”、从“不清晰”到“清晰”的转变。

3) 形成正确决策

数据的作用还在于能让人们发现问题，并形成正确的判断与决策，告诉人们应该做什么、怎么做。只要相信数据的力量，数据就能创造信任与价值，帮助人们作出正确的选择。数据有助于分析人类社会的发展规律和自然界的发展规律，利用大数据分析结果可以归纳和演绎出事物的发展规律，掌握事物的发展规律有助于人们进行科学决策。

4) 保护个人隐私

随着人们全面进入信息时代和数字经济时代，个人数据已成为时刻伴随着人们的个人标签，成为识别个人身份、彰显个人能力、证明个人信用和体现个人行为的重要信息。在互联网经济中，个人数据对于每个人来说也是一种资产，但个人数据的所有权很难进行清晰界定。国家不断出台相应的法律、法规以保障个人数据的安全，并以文件形式多次强调数据的重要性和安全性，如我国出台并施行的《民法典》及其他法律、法规都包含针对个人信息安全的相应规定。数据创造价值的目的是最大化地利用数据，而数据利用的前提和基础是保护个人隐私。随着科学技术的不断发展，通过数据处理、计算、管理等措施可有效防止个人隐私的泄露与滥用。国内外利用隐私计算平衡隐私保护和数据价值流转

已经取得了积极成效。运用隐私计算技术，“去标识化”后的数据可以满足绝大部分个人隐私保护的要求。

2. 数据对于企业的价值

数据已经成为新的生产要素，在数字经济时代拥有巨大的潜力。数据的使用成为提高企业竞争力的关键要素。数据对于企业的价值是由市场需求和经营管理两部分决定的。

从市场需求这一角度来说，数据要素的价值在于重建了对市场理解、预测和控制的新体系、新模式。这种模式本质上是用数据驱动的决策替代经验决策，即基于数据加上算力和算法可以对物理世界进行描述、原因分析、结果预测和科学决策。对于企业来说，数据要素创造价值不在数据本身，数据只有与基于商业实践的算法、模型聚合在一起才能创造价值。数据要素融入劳动、资本、技术等每个单一要素中，能够提高单一要素的生产效率，单一要素的价值会倍增，更重要的是提高了这些传统要素之间的资源配置效率。通俗地说，数据本身生产不了汽车和房子，但是数据要素与各要素聚合有助于低成本、高效率、高质量地生产汽车和房子。数据要素推动了传统生产要素的革命性聚变与裂变，成为驱动经济持续增长的关键因素，这才是数据要素的真正价值所在。数据有助于企业激活其他要素，提高产品研发和商业模式的创新能力，提高个体及组织的创新活力。数据要素的价值在于使企业用更少的物质资源创造更多的物质财富和服务，并对传统生产要素产生替代效应。

从经营管理的层面来看，数据的价值主要体现在记录、备份、监督、纠偏和预测方面。数据本身被记录下来其实是一种操作的基础，脱离了数据记录，后续的操作将难以进行。数据记录对于企业来说更多的是以前操作过程的虚拟备份，记录了不同操作的步骤及次序，使得操作的情景有了复原的可能性，数据记录和备份可以看作数据的“初始价值”。在万物互联互通时代，不同主体之间的交互不断增加，情景复原除了具有纪念意义外，还有一个重要的价值就是事后责任追究，让每个主体对自己的行为承担责任，使各种有效连接成为一种可能。企业备份数据还有更重要的溯源价值，即对于每一操作环节进行有效监督，一旦出现问题就可迅速响应并采取应对措施，具体问题具体分析，以利于事后责任追究，最大限度地降低企业的风险成本。作为一类特定系统，企业在运转过程中需要保证内部各方面的平衡，数据展示出的差异有利于纠偏，防止因打破平衡而造成不必要的损失。伴随着数据技术的不断发展，让数据驱动决策帮助企业降本增效成为各企业的战略目标。数据给予企业的最重要价值之一还包括对未来业务方向的预测作用。企业对所记录的各种数据进行深入比较和研究，发现其中规律特征，依此对系统及业务模式进行优化升级，根据预测结果了解市场长期演变方向，对企业有效制定可持续发展规划意义重大。

3. 数据对于国家的价值

数据作为生产要素的一个重要背景是数字经济的蓬勃发展。数字经济以数据为关键生产要素，以现代信息网络为重要载体，以信息通信技术的有效使用为效率提升和经济结构优化的重要推动力，是以新一代信息技术和产业为依托的新经济形态。数字经济的构成明显要比传统的农业经济和工业经济更加广泛。农业经济体系属于单层结构，以农业为主，配合以其他行业，以人力、畜力和自然力为动力，使用手工工具，以家庭为单位自给自足，社会分工不明显，行业间相对独立；工业经济体系是两层结构，即提供能源动力和行



业制造设备的装备制造产业，以及工业化后的各行各业形成分工合作的工业体系。数字经济体系则可分为3个层次：提供核心动能的信息技术及其装备产业、深度信息化的各行各业以及跨行业数据融合应用的数据增值产业。数字经济是新兴技术和先进生产力的代表，把握数字经济发展大势，以信息化培育新动能，用新动能推动新发展，已经成为全球经济发展的共识。近年来，在以习近平同志为核心的党中央的坚强领导下，我国数字经济已初步建立了顶层引领、横向联动、纵向贯通的战略推进体系，并取得了较好的成绩。据统计，2012—2021年，我国数字经济规模从11.2万亿元增长到45.5万亿元，总量居世界第二，占GDP的比重从20.8%增长到39.8%，已经成为经济高质量发展的关键支撑。

数字经济以数据为关键生产要素，数据正成为与物质资产和人力资本同样重要的基础生产要素。一个国家拥有的数据规模及运用数据资源的能力将成为综合国力的重要体现，对数据的占有权和控制权将成为陆权、海权、空权之外的国家核心权力。数据作为驱动创新发展的关键生产要素，在全球范围引领社会变革，促进透明政府的发展，形成以人为本的数据战略。

数据在全球经济运转中的价值日益凸显，世界主要经济体围绕数据资源抢夺数字经济制高点的竞争日趋激烈。数据价值持续溢出，不仅代表着数据在社会经济发展中的地位不断提升，也标志着数据的含义在不断演变。当前，世界主要国家和地区都已认识到数据对于提升社会经济发展和国家实力的重要意义，通过出台国家数据战略、完善国内数据立法、加强国际数据合作等多种方式，促进本国数据资源开放和数据技术开发。对于一个国家而言，数据已经逐渐渗透到国家社会经济领域中的每一个角落，关乎国家的发展与安全，是重要的国家资产。美国十分重视数据这一国家资产的建设和保护，2012年奥巴马政府发布了《大数据研究和发展倡议》（Big Data Research and Development Initiative, BDRDI），将数据定义为“未来的新石油”；2019年12月美国发布《联邦数据战略与2020年行动计划》（Federal Data Strategy&2020 Action Plan, FDSAP 2020），“将数据作为战略资源进行开发”成为美国新的数据战略的核心目标。欧盟针对数据开放、数据流通、发展数字经济发布了《迈向繁荣的数据驱动型经济》（Towards a Thriving Data-Driven Economy, TTDE）、《建立欧洲数字经济》（Building a European Data Economy, BEDE）、《迈向共同的欧洲数据空间》（Towards a Common European Data Space, TCEDS）等多个战略文件。2020年2月，欧盟委员会发布《欧洲数据战略》（A European Strategy for Data, ESD），强调提升对非个人数据的分析利用能力。2020年9月，英国发布《国家数据战略》（National Data Strategy, NDS），提出释放数据价值是推动数字部门和国家经济增长的关键。

继我国在2019年通过的《决定》中首次提出将“数据”作为生产要素后，中央第一份关于要素市场化配置的文件《关于构建更加完善的要素市场化配置体制机制的意见》（以下简称《意见》）于2020年4月9日发布。《意见》提出了在土地、劳动力、资本、技术、数据5个要素领域改革的方向，明确了完善要素市场化配置的具体举措。数据作为新型生产要素，《意见》为其参与收益分配解除了制度障碍，成为全社会备受关注的焦点。《意见》再一次将数据与传统生产要素相提并论意味着数据要素对于国家的重要程度显著提升。2020年的政府工作报告中强调，要推进要素市场化配置改革，培育技术和数据市

场，激活各类要素潜能。2021年11月30日，工业和信息化部发布《“十四五”大数据产业发展规划》（以下简称《规划》）。《规划》指出，“数据是新时代重要的生产要素，是国家基础性战略资源”。由此可见，数据作为国民经济基础性战略资源的重要地位日益凸显。

1.2 数据治理的内涵

两个人交换自己手中的物品，每个人还是只能拥有一件物品，但若他们交换的是手中的信息，那每个人便得到了两条信息。数据价值的实现也是如此，只有数据流通起来，蕴含其中的价值才能得以发挥。数据治理的本质便是要实现数据的流动，在互联互通中最大限度地挖掘和释放数据的价值，避免数据成为一潭“死水”、一个个“孤岛”。本节主要介绍数据治理的内涵，包括数据治理的概念、数据治理的要素与特点、数据治理的类型与领域、数据治理的内容与范围、数据治理的相关道德准则以及数据治理的目标与意义。数据治理的内涵是数据治理的基础，把握其相关内容有助于理解数据治理的发展，同时能够更好地认识数据的价值。



视频讲解

1.2.1 宏观层面数据治理的基本概念

1. 数据治理的内涵

宏观层面数据治理又称为广义的数据治理，其内涵是对数据资产管理行使权利和控制的集合（包括规划、监控和执行），指导其他数据管理职能如何执行，在高层次上执行数据管理制度。组织为实现数据资产价值最大化所开展的一系列持续工作过程，明确数据相关方的责权、协调数据相关方达成数据利益一致、促进数据相关方采取联合数据行动。

2. 数据治理的核心内容

基于上述概念，我们可以明确数据治理的几个核心内容。

（1）以释放数据价值为目标。数据治理的首要目标是通过系统化、规范化、标准化的流程或措施，促进对数据的深度挖掘和有效利用，从而将数据中隐藏的巨大价值释放出来。

（2）以数据资产地位确立为基础。由于数据治理以数据为对象，那么作为核心要素，数据在社会经济发展中所处的地位直接决定了围绕数据的各项活动的开展方式、流程等。

（3）以数据管理体制机制为核心。数据治理的重点在于建立健全规则体系，形成多方参与者良性互动、共建共享共治的数据流通模式，因此，围绕数据的各项管理体制机制的建立和完善是当前国家、组织、企业等各类主体的核心。

（4）以数据共享开放利用为重点。数据治理的目标在于保障数据的有序流通，进而不断释放数据的价值。而数据流通的主要活动包括数据的共享、开放以及有序的开发利用等，这也成为当前阶段数据治理工作的重点。

（5）以数据安全性与隐私保护为底线。数据治理要以国家、企业和个人信息安全为前提，否则再好的数据治理模式也是有违社会正义的。因此，保障数据安全性与隐私保护的各项活动是数据治理的底线保障。



3. 数据治理的核心目标与价值取向

综合来说，数据治理的核心目标就是通过各种方式提升数据的价值，而提升数据价值的

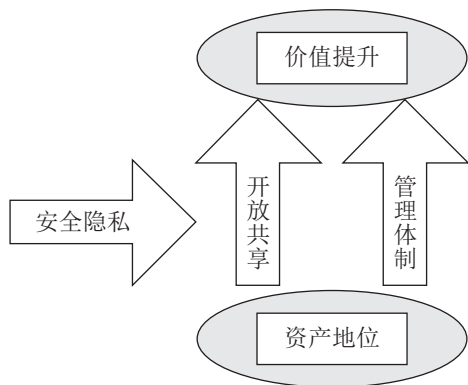


图 1.2 数据治理的核心目标

的核心就是确定数据的资产地位，如图 1.2 所示。为了提升数据的价值，需要系统地设计管理体制机制，包括数据治理组织和数据管理活动；需要最大限度地推动数据开放共享，没有数据的开放共享就没有数字经济的发展。当然，这一切需要有数据安全和隐私保护的底线作为保障，否则数据的价值就难以得到体现。

数据治理的目标也就是其所需要达成的目的和结果。总体来说，就是以治理的力量，在数据管理的基础上，进一步优化数据价值实现的关键和核心因素，确保实现高效能的数据功效的倍增。

从数据治理的现实出发，数据治理的目标应当具备以下 4 方面的特性：合法合规性、科学可行性、高度契合性、动态适应性。合法合规性，即数据治理的目的和结果都必须遵守国家政策和法律法规，与党和国家的根本意志和公共利益永远保持高度契合；科学可行性是指数据治理的目的和结果必须是符合客观规律的，是可以被信赖和依靠的，是科学合理、经过努力就能实现的；高度契合性，即数据治理的目的和结果必须与国家经济社会发展的要求保持高水平的匹配和适合，特别是要与国家经济社会发展资源结构优化的客观要求保持一致；动态适应性表明数据治理的目的和结果不是一成不变的，需要根据不断变化的经济社会发展的需要和要求，不断进行动态的适应性调整。

综合上述 4 个特性，可以提炼出数据治理的目标，即根据国家经济社会发展资源结构调整的需求，遵循数据价值实现的规律，构建有利于数据开发利用与管理的体制机制和环境氛围，强化监督控制和公共服务，强化社会参与和深度合作，促进和保障数据产业健康、持续发展，切实维护公民个人和社会组织的合法权益，确保实现数据资源功效的倍增，全面支持经济社会的可持续发展。这一目标的具体内涵和要求同样包含 4 个方面：

(1) 制定数据治理目标的基本依据是国家经济社会发展资源结构调整的客观需求，对数据实施治理的目的就是满足这种关乎国家经济社会发展全局和根本的客观需要和要求。

(2) 对数据进行治理的结果就是使数据资源得到有效的开发、利用和管理，数据产业获得健康、持续发展，公民和各种社会组织的合法权益得到切实维护，确保实现数据资源功效的倍增。

(3) 对数据实施治理的客观依据是数据资源价值实现的规律，这是数据开发、利用和管理能够真正纳入科学轨道、取得实实在在的效力和功用的保障。

(4) 实现数据治理目标的基本途径是构建有利于数据资源开发、利用与管理的体制机制和社会环境，强化监督控制和公共服务，强化广泛的社会参与及深度合作。

数据治理的基本价值取向是数据治理所依据的价值尺度，体现了治理的核心主张，反映了治理主体的目标、动机、基本态度和决策意图。价值取向是治理内容中最重要的部分，结合经济社会发展的现实情况，现阶段数据治理的基本价值取向应该包括：

- (1) 数据治理的目的和结果具有多元性特征。
- (2) 数据治理的重点依据治理层级的不同而有所区别。
- (3) 数据治理需要强有力的科学组织，加强对数据资源的规划和配置。
- (4) 数据治理以优化的制度建设和强有力的制度实施为基础和特征。
- (5) 大力推动和鼓励开发利用数据资源。
- (6) 有力维护国家数据主权，尊重和保护各种主体合法的数据资源权利。
- (7) 明确数据资源所有权的形式，依法处置与数据相关的权属体系。
- (8) 明确规定不同所有权的行使规则。
- (9) 加强包括数据隐私和数据安全在内的数据资源保护体系的建设。
- (10) 加强数据资源的建设，大力促进和保障数据产业的发展。

4. 数据治理的特点

数据治理具有以下特点。

(1) 以人和数据为中心，强调满足多元主体的价值追求和利益诉求。从数据治理的定义可以得出，数据治理要实现数据在全生命周期过程中安全、有效流通所执行的操作。在这一过程中数据是操作的对象，人是操作的主体，而数据治理成果的获益者包括国家、企业和个人。因此，构建数据治理的体系也要考虑这3个层面。国家、企业和个人所产生和使用的数据的形式不尽相同，因此要注重数据治理过程中满足多元主体的价值追求。如在国家层面，通过制定“上位法”，明确数据的权属和合理使用数据的边界；通过成立国家标准化管理委员会等多级机构，领导数据治理相关的标准工作；在司法领域和政府数据开放两方面，也有不少应用实践的案例；通过科技部、国家自然科学基金委员会等部门，组织与数据治理有关的科研项目，引导数据治理的支撑技术研究。在企业层面，要在遵守国家规范的基础上建立数据治理的专职机构，协调各业务部门并确保数据治理在整个组织内得到支持。在个人层面，要注重个人隐私的保护以及数据权属的界定等问题。

(2) 从数据安全和数据主权的高度出发，专注于数据的全生命周期安全。数据安全治理体系的建设要以数据全生命周期为核心，实现数据安全全方位治理。传统的数据安全监管方法以系统为中心，但目前，数据的共享交换已经变成同一个部门、不同层级之间流动的常态化过程，所以构建数据全生命周期的监管体系势在必行。数据安全治理体系框架通过3个维度构建而成，包括政策法规、技术层面和安全组织人员。数据安全治理体系框架的建设应在符合政策法规及标准规范的同时，在技术上实现对数据的实时监管，并配合对安全组织人员的规范培训。在整个体系中，核心监管技术体现在技术架构层面，包括安全运营中心、数据中心以及安全基础资源。通过提供最基础的技术保障，使安全运营中心对整个数据中心进行实时的响应控制。安全运营中心的作用集中体现在资产管理、合规监管、实时监测、数据安全态势以及通报预警等方面。安全运营中心通过采集数据中心数据，进行数据汇聚、分析以及治理从而实现对数据的实时管控。数据安全基



础资源是整体技术框架的支持组件，在提供最基础的技术保障的同时，以工具的形式保障数据安全。

(3) 强调数据资源，尊重数据规律，遵循数据伦理。在数据生命周期的每个阶段，都必须始终考虑道德伦理和隐私问题，从保护个人数据的隐私收集到基于自动数据分析进行的决策，数据生命周期中的伦理探究始终是需要不断关注的方向。数据伦理问题的探析与数据生命周期密不可分，但目前数据伦理问题的研究和解决方案大多只针对生命周期的某一阶段，而隐私伦理问题和安全威胁存在于整个数据生命周期中，应从各个阶段详细探析。总体上看，数据生命周期的有序递进，离不开隐私安全和伦理道德规范的规制。应当从数据生命周期视角研究伦理问题，明晰数据在不同阶段下的状态，有效识别杂糅在各个阶段的伦理问题及潜在风险，以确保数据使用主体更恰当地使用数据，促进数据应用符合法律与道德标准。

(4) 与具体情景治理理念和组织业务目标保持一致。离开现实背景的支持，再多的理论都只是无源之水、无本之木，因此数据治理的一个重要特点是与情景治理理念和组织业务目标保持一致，在此基础上，数据治理才能最大限度地挖掘数据要素的价值，并对组织目标的实现起到积极作用。正因为数据治理过程对组织目标的推进作用，因此在组织内的数据治理实施绝不是一个部门的事情，需要从组织的全局性视角考虑，建立专业清晰的数据治理组织架构，明确权责关系，培养整个组织的数据治理意识，从而保障数据治理过程中数据的质量和治理的效率，从而为关键业务和管理决策提供支持。从总体上说，数据治理组织体系应如图 1.3 所示。

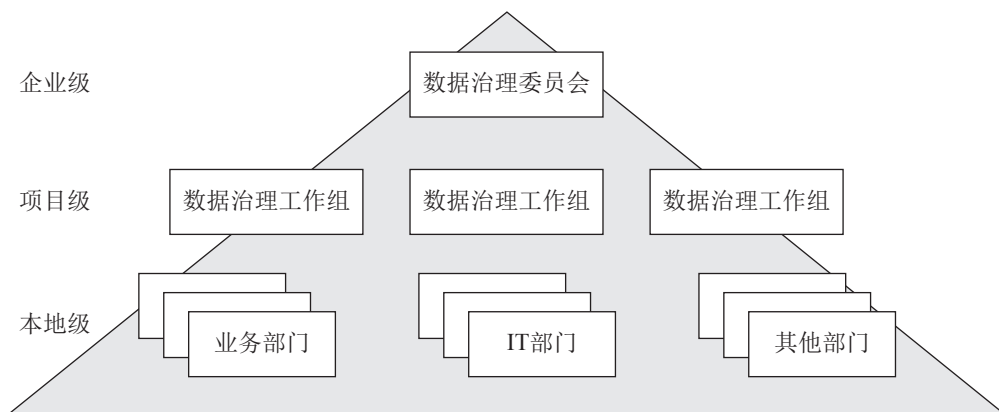


图 1.3 数据治理组织体系示意图

1.2.2 数据治理工作的定义、要素、类型和领域

1. 宏观数据治理与微观数据治理的区别

宏观数据治理是国家层面的数据治理，也是国家政策和法律法规中经常出现的数据治理一词的含义，属于国家治理范畴。微观数据治理是企业层面的数据治理，也可称为企业数据治理或者数据治理工作，属于企业管理的范畴。本书将微观数据治理称为数据治理工作（在第 8 章则称为企业数据治理），或者在不存在混淆可能性的情况下直接称为数据治理。

2. 国际机构对微观层面数据治理的定义

纵观国际各个机构对数据治理的研究，其对数据治理的定义存在差异。

(1) 国际电工委员会 (International Electrotechnical Commission, IEC) 在 IS/IEC TR 38505-2:2018 (Information technology. Governance of IT. Governance of data-Implications of ISO/IEC 38505-1 for data management, 信息技术 IT 治理 数据治理 第 2 部分: ISO/IEC 38505-1 对数据管理的影响) 中对数据治理的定义。数据治理是关于数据采集、存储、利用、分发、销毁过程的活动的集合。

(2) 全国信息技术标准化技术委员会在 GB/T 34960.5—2018 (Information technology service-Governance-Part 5: Specification of data governance, 信息技术服务 治理 第 5 部分: 数据治理规范) 中对数据治理的定义。数据治理是数据资源及其在应用过程中相关管控活动、绩效和风险管理的集合。

(3) 国际数据管理协会 (Data Management International, DAMA) 对数据治理的定义。数据治理是指对数据资产管理活动行使权利和控制的活动集合 (规划、监督和执行)。

(4) 国际数据治理研究所 (Data Governance Institute, DGI) 对数据治理的定义。数据治理是一个通过一系列信息相关的过程来实现决策权和职责分工的系统, 这些过程按照达成共识的模型来执行, 该模型描述了谁能根据什么信息, 在什么时间和情况下, 用什么方法, 采取什么行动。

综合分析以上各机构对数据治理的定义可以看出, 微观数据治理又称为狭义数据治理, 是指为确保数据安全、私有、准确、可用和易用所执行的所有操作, 包括主体必须遵循的流程、需要采取的行动以及在整个数据生命周期中为其提供支持的技术。数据治理是在数据生命周期 (从获取、使用到处理) 内对其进行管理的原则性方法。数据治理意味着制定适用于收集、存储、处理数据的内部标准, 即数据策略, 对数据的访问权限、修改权限等做出统一的规定。此外, 数据治理还涉及行业协会、政府机构及其他主体设定的外部标准。

数据治理以数据为对象, 由于数据的来源、流通具有高度复杂性, 因此数据治理是一个复杂的过程, 包括数据采集、归集存储、分析处理、数据产品和服务定价与分配等多个复杂的流通环节; 涉及数据生产者、数据采集者、数据管理者、数据平台运营者、数据加工利用者等多元参与主体 (政府、市场、社会), 是一个复杂的动态变化过程。

3. 数据治理工作要素

数据治理是一个系统工程, 是一个从上至下指导, 自下而上推进的工作。因此, 在指导方面必须得到大家的共识, 要有强有力的组织、合理的章程、明确的流程、健壮的系统, 这样才能使数据治理工作得到有效的保障。数据治理工作的要素包含以下几点。

1) 发展战略目标

战略是选择和决策的集合, 二者共同制定出一个高层次的行动方案, 以实现更高层次的目标。数据战略是每一个组织发展战略中的重要组成部分, 是保持和提高数据质量、完整性、安全性的计划, 是指导完成数据治理工作的最高原则。

2) 数据治理组织

数据治理的组织包括制度组织和服务组织。制度组织主要负责数据治理和数据管理制



度。制度组织是跨职能部门的组织，负责解决整体数据战略、数据政策、数据管理度量指标等数据治理规程问题。服务组织根据数据治理领导小组制定的愿景和长期目标，协助各数据领域工作层级实施数据治理工作，对整体数据治理工作进行度量、检查和汇报，并对跨领域数据治理问题和争议进行裁决和统筹决策。

3) 制度章程

制度章程是确保对数据治理进行有效实施的责任制度，包括数据治理职能的职责，也包括其他数据管理职能的职责。数据治理是最高层次的、规划性的数据管理制度活动，是主要由数据管理人员和协调人员共同制定的高层次的数据管理制度决策。数据治理的制度章程一般包含规章制度、管控办法、考核机制与技术规范。

4) 流程管理

流程管理包括流程目标、流程任务、流程分级。根据数据治理的内容，建立相应的流程，且遵循这一组织数据治理的规章制度。实际操作中可结合所使用的数据治理工具，与数据治理工具供应商协商，建立符合组织条件的流程管理。

5) 技术应用

技术应用包括支撑核心领域的工具和平台，例如，数据质量管理体系、元数据管理系统等。这是数据治理能够顺利开展的技术保障。只有建立丰富的数据治理工具和平台，才能在各个领域有效地进行数据的管理和治理，才能有效提高数据价值。

6) 监督评估

监督评估是数据治理中不可缺少的要素。数据治理离不开监督，一旦在治理过程中破坏了数据的完整性或一致性，通过监督溯源就可及时发现问题所在。一个组织的数据治理效果是否达到预期，可以通过数据的质量、数据治理的效率、数据治理模型的成熟度几方面进行评估。

4. 数据治理工作的类型

数据治理工作主要分为以下 3 种类型。

1) 基础数据治理

基础数据用于对其他数据进行分类，因此也称作参考数据。基础数据通常是静态的（如国家币种、个人信息等），一般在业务事件发生之前就已经预先定义。基础数据的可选值范围有限。当基础数据的取值发生变化的时候，通常需要对流程和信息系统进行分析和修改，以满足组织业务需求。因此，基础数据的管理重点在于变更管理和统一标准管控。

2) 主数据治理

主数据是参与业务事件的主体或资源，是具有高业务价值的、跨流程和跨系统重复使用的数据。主数据与基础数据有一定的相似性，都是在业务事件发生之前预先定义；但主数据又与基础数据不同，主数据的取值不受限于预先定义的数据范围，而且主数据记录的增加和减少一般不会影响流程和系统的变化。

3) 事务数据治理

事务数据在业务和流程中产生，是业务事件的记录，其本身就是业务运作的一部分。事务数据是具有较强的时效性的一次性业务事件，通常在事件结束后不再更新。

1.2.3 数据治理工作的内容与范围

1. 数据治理工作的内容

数据治理工作的内容是指数据治理的落脚点。国际数据管理协会指出，数据治理的主要内容包括数据架构管理、数据质量管理、元数据管理、文档与内容管理、数据仓库与商业智能管理、主数据管理、数据安全管理等。国际数据治理研究所指出，数据治理核心领域包括：政策、标准和策略；数据质量；隐私、合规和安全；架构和集成；数据仓库和商业智能；管理支持。全国信息技术标准化技术委员会指出，数据治理的内容包括数据管理体系和数据价值体系，其中数据管理体系包括数据标准、数据质量、数据安全、元数据管理和数据生存周期，数据价值体系包括数据流通、数据服务、数据洞察等。尽管不同领域的数据治理内容各不相同，但从当前学界研究和实践应用的共识来看，数据治理的主要内容需含有元数据管理、主数据管理、数据质量管理、数据安全管理的 4 个基础方面。

2. 数据治理工作的范围

数据治理工作是在国际协作、国家治理、行业监督和企业管理中，为了提升数据的质量、降低数据管理成本、保障数据安全和管控数据风险，针对公共数据、政府数据、企业数据和个人数据的采集、存储、应用和流通等一系列环节，利用各类工具方法进行有效管理，主要包括法律法规、行业标准、企业制度、技术工具等。为明确数据治理的范围，利用数据和治理“4W1H”模型进行说明，如图 1.4 所示。遵循数据治理的概念内涵及标准化的自身含义，数据治理为政府、企业、个人提供服务，保障各类数据全生命周期的有序运转。

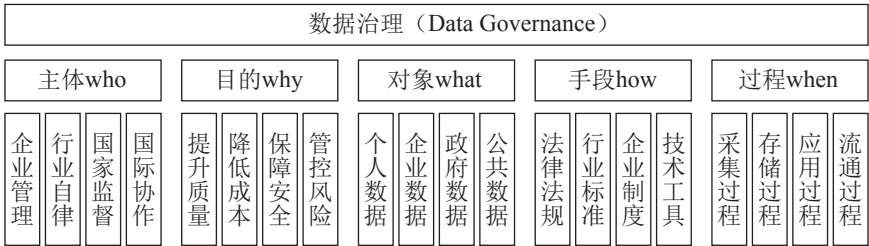


图 1.4 数据治理的范围

1.2.4 数据治理工作的目标与意义

1. 数据治理工作的目标

数据治理的目标是对组织的数据管理和利用进行评估、指导和监督，通过提供不断创新的数据服务，提高数据利用效率，为组织创造价值。国际数据管理协会指出，数据治理的目标包括：落实数据政策、标准、计划、指南等；跟踪数据政策、标准和规则的遵从；解决数据质量问题；提升数据价值。国际数据治理研究所指出，数据治理的通用目标包括更好地决策、减少运营摩擦、保护数据相关者的权益、建立标准化及可重复的流程、降低成本并提高效率、确保过程透明等。

2. 数据治理工作的意义

高质量的数据对任何企业都是战略性资产，随着企业推进数字化转型的进程，有效数



据正迅速成为一个关键的业务差异，但要使数据具有价值，需确保数据的可信性、安全性、可访问性、准确性、共享性和及时性。数据治理有助于增强企业灵活性，实现相关成本和风险的最小化决策，特别在数字经济中，数据治理比以往任何时候都显得更为重要。

缺乏企业高层领导的支持、系统间的数据壁垒和整个治理项目缺乏明确的流程和数据模板、数据所有权和问责机制不清等是导致项目失败或者治理结果不理想的主要原因，涉及企业中所有跨功能和跨业务的决策机制。数据治理具有战略性、长期性、艰巨性、系统性、持续企业内部数据环境优化治理工作，因此数据治理是一个漫长而持续的过程，没有一针顶破天的诀窍，也没有立竿见影的捷径，要避免对数据治理工作粗浅的认识。

数据治理工作的意义在于更加有效、合规地使用数据，以及与业务相关的各种数据的不断增加，是推动数据治理的主要动力。数据治理要了解各级政府和企业对数据的不同需求，并按自上而下的方针，对数据进行调研和管理，全面理清数据资产的分布，对数据的管理、应用、质量等各方面进行全面科学的评估，从而更好地对数据产生、采集、处理、加工、使用等过程进行规范。数据治理就是要对数据统一标准，出台合理的治理流程和制度，规范各类数据的生产供应。只有通过数据治理，不断提升数据质量，严格控制数据安全，才能让数据在数字化转型中发挥出最大的效益。

1.2.5 数据治理的相关道德准则

1. 保护隐私安全

保护隐私安全这一准则意为为数据治理中个人、组织及国家的合法数据权益提供基础保障。数据治理的开展应当合法、正当地保护个人隐私与数据安全，不得以任何违反个人意愿及国家规定的方式非法收集利用相关数据。

2. 促进公平公正

促进公平公正是数据治理道德性实现的基础准则。这一准则对于各相关主体的合法权益切实进行保护，有助于促进社会公平正义和机会均等，可以有效预防、杜绝垄断市场等恶性事件的发生，推动全社会公平共享数据治理带来的成果。

3. 增强透明可信

增强透明可信的道德准则旨在打破技术壁垒所带来的道德失范。“透明可信”并非指“全面的、简单的、绝对的公开”，而是指在数据治理与产业激励之间寻求信任与合作的平衡。此外，增强透明可信的准则对于数据治理的可溯源性提供了保障，与数据治理的安全可控准则相互包容。

4. 发挥数据要素作用

发挥数据要素作用的准则在于通过数据治理加快数字经济、数字社会的建设，其与国家基础设施建设及综合能力提升相协调。数据作为数据治理的基础，在社会生产活动中发挥着越来越重要的作用。随着数据要素市场培育的加快，数据治理能力不断提升，带动社会生产和治理方式的转变，全面推进数字化国家建设。数据要素市场培育的发展方向如图 1.5 所示。

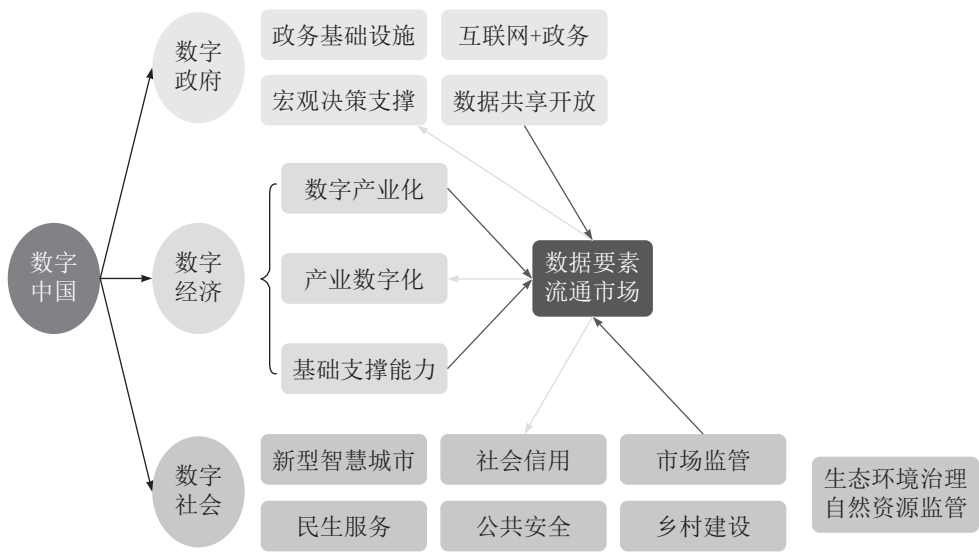


图 1.5 数据要素市场培育方向

1.3 数据治理的研究进展

为了解决数据流动过程中产生的一系列问题，“数据治理”一词逐渐兴起，业界学界的研究都在如火如荼地开展，本节主要介绍数据治理发展至今的探索研究成果，包括理论发展、内容研究、框架体系、应用实践等，建立对数据治理研究进程的概论性了解。

1.3.1 数据治理的研究起源

数据治理的概念最早源于企业，2004 年，H. Watson 探讨了“数据仓库治理”在公司中的实践，此后“数据治理”在企业管理中的研究不断得到关注。随着数据治理研究的深入，涌现出了一系列代表性的成果，这就是数据治理的研究源起。随着现实中数据治理探索的发展，数据治理的研究过程和方法都发生了显著的变化。

首先是数据治理的核心概念及其关系研究。数据治理的概念与数据资产管理、IT 治理、数据管理等概念都有一定的密切联系，实际上，这也侧面反映了数据治理研究的不同视角。

数据资产管理的内涵是控制和保护数据，发掘数据资产的潜在价值。在宏观规划视角，进行数据资产管理战略计划的设计、实施、监控。从中观视角看，实际上是对数据资产管理计划的整体活动过程的管理。

IT 治理的内涵是在公司治理或组织治理中，进行责任构建、战略规划、数据接收、协调管理、组织行为等方面的活动，属于中观视角的数据治理。

数据管理的内涵是进行数据过程管理和数据质量管理。这是从微观的实施视角去进行数据治理，要求数据具有一致性、可用性、安全性，涉及数据所有权和数据管理责任问题，属于数据治理的微观实施范畴。

数据治理的内涵是对组织的数据管理和利用进行评估、指导和监督，以及通过不断提



供创新的数据服务的方式，为组织创造价值。

数据治理与数据管理是不同的活动。数据治理负责对管理活动进行指导、监督与评估，而数据管理则是根据所作出的决策来具体计划、建设和运营。正如习近平总书记所指出的“治理和管理一字之差，体现的是系统治理、依法治理、源头治理、综合施策”。数据治理是联合行动的过程，强调协调而不是控制，本意是通过服务来达到管理的目的。数据治理更多的是面对战略层面、组织层面、制度层面的事务，确立“什么样的决策需要在什么层级制定”。数据管理是操作和实施层面的概念，是通过一系列实际落地的办法去实现“治理”目标的具体过程。数据管理是通过建立组织结构和工作机制，落实谁牵头、谁配合、谁主责、谁落实，在各自的职能领域去完成具体任务，包括企业级层面的数据标准化、数据资产管理，业务领域层面的数据规范化、数据质量改进等。数据管理所涉及的领域范围相对数据治理来说更加细节化和行动化，数据治理则更意味着为实现数据价值的一系列制度安排和举措。数据治理核心概念的辨析如表 1.2 所示。

表 1.2 数据治理核心概念的辨析

核心概念	发布时间	目标	来源
数据资产管理	2014 年	①组织数据资产审计、登记 ②控制并保护数据，发掘数字资产的潜在价值，提高数据利用效率	《资产管理 综述、原则和术语》(ISO 55000:2014)
IT 治理	2015 年	①公司治理和广义的组织治理 ②责任构建、战略规划、数据接收、绩效管理、协同管理、组织行为	《信息技术 组合 IT 治理》(ISO/IEC 38500:2015)
数据管理	2019 年	①数据过程管理 ②数据质量管理	《技术规范 D2.1 物联网和智慧城市及社区数据处理和管理框架》(ITU-TFG-DPM D2.1)
数据治理	2020 年	①对组织的数据管理和利用进行评估、指导和监督 ②通过提供不断创新的数据服务，为组织创造价值	《信息技术 大数据参考架构 第三部分：参考架构》(ISO/IEC 20547-3:2020)

对数据治理模型的理论研究也取得了一定的成果。数据治理模型可以帮助组织厘清复杂、模糊的概念及关系，指导组织开展高效的数据治理工作。当前关于数据治理模型的研究主要集中在数据治理成熟度评估、数据治理有效性评价、面向大数据背景下的数据治理模型和针对具体应用层面的数据治理模型研究几方面。

此外，还有数据治理框架的相关理论研究。数据治理框架是指基于对数据治理领域基本概念的分析，建构起相关的概念关系，以反映这一领域中的基本概念及概念间的逻辑关系，从而为实现数据治理战略目标提供理论基础。当前国内对于数据治理框架的研究主要集中在国际主流的数据治理框架、大数据情境下的数据治理框架和针对具体行业领域的数字治理框架构建等几方面。

1.3.2 数据治理的研究重点

随着各类组织中业务的增长，海量、多源异构的数据对数据的管理、存储和应用均提

出了新的要求。因此，顺应时代发展趋势，构建起完整的数据治理体系，提供全面的数据治理保障，从而充分发挥数据的资产价值，更好地支持数据治理的应用实践，成为学术界关注的焦点问题。

当前应用研究的主要内容可以概括为以下几方面：

(1) 数据治理的体系构建研究，包括数据治理模型与框架的设计与验证等。

(2) 数据治理的有效保障机制研究，就如何提升数据质量、保护数据隐私、保证数据安全等关键问题进行深入探讨。

(3) 数据治理在具体应用领域的实践探索，特别是以数据为核心的行业的数据治理问题。

(4) 大数据、全球化时代背景给数据治理带来的新机遇与挑战。

当前，国内关于数据治理的研究基本是偏重理论，大多是侧重数据治理的规划、方法、框架、体系构建等，数据治理技术层面的研究较少。

1.3.3 数据治理的应用场景

数据治理发源于图书馆、电力、银行业等以数据为核心业务的行业，随着数据的爆发式增长，高校建设、医疗层面和政府治理等方面的数据治理问题也得到了学界的重视。数据治理应用实践层面的研究从不同的行业，针对不同领域所遇到的实际问题展开。通过将数据作为组织的资产进行管理，运用有效的数据治理手段，发挥数据的价值，进而提升组织的竞争力。数据治理的场景研究对于未来的经济增长和社会发展均具有重大意义。

下面给出数据治理的一些典型应用场景。

(1) 政府数据治理领域。基于我国政府部门间的数据协同、数据治理形态、数据利用和数据流程等问题进行研究，分析我国数据治理能力的现状和不足，总结可借鉴的国际经验，提出我国政府部门数据治理的实践路径。

(2) 图书馆数据治理领域。以图书馆数据为研究对象，结合高校图书馆数据的类型和特点，基于国际上权威的数据治理框架提出我国高校图书馆的治理框架，阐明大数据时代图书馆职能的转变。

(3) 医疗领域。分析医疗健康大数据资源特性及治理问题，探索医疗健康数据治理的实施步骤，包括组织、执行、监管等方面，建立医疗健康数据治理体系。

(4) 金融领域。探索数据标准化、数据治理规范化运作流程，赋能产业链数据管理，提升监管数据在源系统等环节的数据治理质量，加速监管数据质量能力体系的搭建进程。

1.4 数据治理的发展

本节介绍数据治理的发展，主要包含数据治理的发展现状、数据治理面临的问题、数据治理的发展机遇以及数据治理的发展趋势。了解数据治理的发展现状可以对数据治理的框架体系有初步的认识，通过国外的数据治理实践总结适用于我国的数据治理经验，对数据治理当前面临的问题提出更好的解决方案。从数据治理的问题与发展机遇中归纳数据治



理的发展趋势，结合数据治理的发展现状，可以更深入地理解数据生产要素的价值维度、价值定位和价值取向。

1.4.1 数据治理的发展现状

数据治理是释放数据要素价值、推动数据要素市场发展的前提与基础。经过多年发展，我国数据治理在政务、金融、通信、电力、互联网等领域已经逐步深化落地。进入2022年，数据治理领域面临的新挑战与老问题共同推动着数据治理实践向前发展。

数据治理一直是国内外研究的热点与重点，数据治理已成为网络空间国际治理领域对话博弈的核心命题之一。国际数据治理的话语对象逐渐从个人数据延伸到非个人数据，这昭示着世界各国围绕数据的合作与竞争正持续深化。目前以欧美为代表的国际数据治理进程已然呈现从立法执法到国际博弈的多层次新走向。着眼于欧盟的数据治理思路，作为对欧盟1995年《个人数据保护指令》（Data Protection Directive, DPD）规范的升级，《通用数据保护条例》（General Data Protection Regulation, GDPR）在新一代欧盟数据治理规范发展历程中具有里程碑意义，其规范约束更深刻地介入数据治理的全生命周期，进而对于大数据、云计算以及人工智能等以数据收集、处理为核心要素的新技术应用产生显著的导向意义，同时也对全球数据治理的未来走向产生了现实的深刻影响。英国历届政府和议会自20世纪90年代起，颁布出台了大量的法律、法规和行政命令，逐步形成了一套相对完整的数据治理政策体系，其内容涉及个人数据（隐私）保护、信息公开（自由）、政府数据开放、国家信息基础设施、信息资源管理与再利用、电子政务和网络信息安全等方面，比如，2018年英国议会通过的新版《数据保护法》（Data Protection Act 2018, DPA 2018）、2000年颁布的《信息自由法》（Freedom of Information Act, FOIA）和2012年颁布的《自由保护法》（Protection of Freedom Act, PFA）等。美国在数据治理进程中不仅出台了具有代表性的法治政策，如《加利福尼亚州消费者隐私保护法》（California Consumer Privacy Act of 2018, CCPA），而且以数据治理为支点驱动数字化城市发展，将纽约市打造成为世界级数据平台，不断扩大数据开放程度，创新开放手段，鼓励公众参与城市数据治理过程，为城市提供优质的公共服务。

我国官方首次正式提出“数据”的概念是在2014年政府工作报告中。此后，其重要性为各级政府所关注。近年来，国家、相关部门和地方政府在数据治理的政策制定、机构设置和专项行动方面取得了显著成效。我国的数据治理政策主要围绕数据基础设施建设、开放共享、示范应用、要素市场、安全保障等方面进行顶层设计，自2018年省级机构改革以来，各地纷纷以不同方式设立或调整合并数据治理机构，形成了以中央主管部门牵头、行业部门专业管理的组织架构。

案例 1.1

江西省农村信用社联合社数据治理实践

2022年，人民银行印发《金融科技发展规划（2022—2025年）》，银保监会发布《关于银行业保险业数字化转型的指导意见》，金融数字化发展迈入“积厚成势”全新阶段。

为解决数字化转型发展瓶颈，构建高质量发展的新动力源，从2019年开始，江西省

农村信用社联合社以“夯实数据基础，提升数据质量，深化数据应用”为目标，持续实施数据治理项目。

“兵马未动，架构先行”，江西省农村信用社联合社数据治理采取了“1+X联邦制”的组织架构模式，以江西省农村信用社联合社领导及相关部门负责人组成的数据治理委员会作为决策管理层，负责各数据治理领域重大事项的审议及监督评价；在委员会下设立数据治理办公室作为组织协调层，负责牵头数据治理各领域工作的开展；江西省农村信用社联合社相关部门和辖内各成员行作为数据治理的工作执行层，设立专职的数据治理综合岗，负责数据治理各项工作的具体落实。通过建立自上而下、协调一致的数据治理体系，明确相关方工作职责，为数据治理有序推进提供强大的组织保障。

项目周期内，江西省农村信用社联合社在全行范围内建立了自上而下的数据治理组织架构，制定了一系列的数据管理相关制度，为数据治理各项活动提供了强有力的组织保障；制定数据标准，夯实数据基础；开展专项治理，提升数据质量；搭建数据管控平台，提升数据治理工作自动化程度。同时，在数据应用方面积极推进，提升了监管报送自动取数率，并以零售业务管理平台为抓手，释放数据治理价值，取得了一定成效，实现了江西省农村信用社联合社数据治理从无到有、从有到优的转变，为推进数字化转型打下了坚实基础。

“工欲善其事，必先利其器。”项目期内，江西省农村信用社联合社搭建了数据管控平台作为数据治理工具，实现了对元数据、数据标准、数据质量的全方位管理，具有元数据血缘分析、标准落标检查、数据质量检核等功能，可以将数据治理中的多种任务线上化、流程化，减少了繁复人工操作，极大地便利了数据治理相关领域的各种类型任务。同时，为充分释放数据治理价值，建设了零售业务管理平台，搭建了数据集市，形成了客户标签体系，构建了大数据分析模型，赋能经营管理和业务拓展，打造互联网时代零售服务内生增力，充分发挥了数据价值。

案例思考题：



>>>

1. 江西省农村信用社联合社开展了哪些具体的数据治理工作？
2. 结合当前数据治理发展现状，江西省农村信用社联合社数据治理工作有哪些值得借鉴的地方？

1.4.2 数据治理当前面临的问题

由于数据治理的相关理论体系尚未健全，因此数据治理在开展过程中遇到了许多挑战与阻碍，数据治理面对着较多的问题。具体体现在如下几方面。

1. 数据治理体系层次不清

根据主体的差异性，数据治理可分为多个层次，不同层次的数据治理有不同的目标和作用对象，其治理措施、范围和效果都有所不同。当前，数据治理主体过于扁平化，在多元共治机制上面临一定的缺失困境，这就使得数据治理体系的层次不清、概念混乱，各级数据治理难以形成统一的整体，作用效果无法达到预期。



2. 数据主权保护权责模糊

数据作为资源蕴涵着宝贵的能量，在当今社会已成为与土地、资本、人力等同样重要的战略资源，数据治理能力也成为国家治理能力的衡量尺度之一。全球范围的云计算和数据跨境流动对数据主权保护提出了新挑战，数据主权的保护成为各级数据治理主体必须考虑的问题。当前，数据治理主体之间的关系并没有明确界定，国际上也没有权威的组织或标准对数据主权的范围进行限定和规范，数据主权保护的责任主体模糊，保护力度有待加强。

3. 数据要素价值实现困难

数据治理的最终目标是实现数据生产要素价值，主要手段是各级数据价值应用。数据治理体系的缺失，使得各级数据应用的目标、策略相对来说都具有一定的局限性，各级主体的数据应用缺乏对数据治理整体的关注与衔接，大多形成了数据孤岛的局面，而缺少交互的数据无法真正发挥数据在要素市场的关键作用。更进一步说，明确数据资产的地位是促进数据要素价值释放的关键。目前，数据的资产地位尚且不明晰，负责数据运营管理的多为成本部门。到目前为止，明确数据资产地位的“上位法”尚未制定。在现行的法律框架下，数据资产无法体现在会计报表上，数据确权、价值衡量、收益分配等问题制约着数据流通以及与之相配套的秩序规则的建立完善，从而影响了数据价值的释放。

4. 数据安全隐私存在风险

在数据主导的时代，安全和隐私是公认的难题。高额收益、低价成本、数据多元异构等不协调因素都增大了数据保护的难度。举例来说，数据隐私保护技术大多基于静态数据集，数据的动态利用使得数据流动过程中的权责难以分辨，数据安全问题更是难以追责。在个人隐私方面，个人数据和数据流动的广泛性增加了个人隐私数据的风险，个人数据权利受到侵害，数据主体对数据的控制权被削弱。尽管当前各级数据治理主体普遍开展了数据安全管理和个人隐私保护的行动，但从整体来看，其策略、措施都存在较大差异甚至存在一定的矛盾点，数据安全与个人隐私保护依旧存在不容忽视的风险。

5. 人才培养体系建设落后

大数据技术与传统产业的深度融合能够促进我国经济发展的新旧动能转换和数字经济的发展，而这一切的发展都需要大量的人才支撑。目前我国数字产业化和产业数字化蓬勃发展，对大数据复合型人才提出了更高的要求：既要懂大数据技术又要懂相关产业的知识。2017年，清华大学经管学院互联网发展与治理研究中心联合全球职场社交平台 LinkedIn（领英）发布的《中国经济的数字化转型：人才与就业》报告显示，大数据与人工智能领域缺口明显，“技术+管理”的复合型人才一将难求。为了适应时代发展的要求，我国已经有众多高校获批数据科学与大数据技术专业，在数据复合型人才培养方面取得了一定的成绩。但在目前的教育体系中，高校人才培养多偏向于理论与技术，与现实业务对接较少。在大数据与各领域融合发展过程是业务导向而不是技术导向，不能将技术和工具限定好再考虑业务，而是要通过业务活动解决实际问题，将技术作为解决问题的工具。

1.4.3 数据治理的发展机遇

尽管数据治理面对的是各类有待解决的问题，但随着近年来新技术及平台经济的发展热潮，数据治理同样也面临着许多选择与机遇。

1. 新技术与数据治理

伴随着大数据、云计算以及机器学习算法的发展，人工智能、区块链等新技术的浪潮从几年前一直延续至今，并且广泛应用于多个行业和领域，成为下一次科技革命的领军技术。同样，伴随着数据量与数据来源的迅猛增长，数据治理也成为了充分挖掘利用数据价值过程中必不可少的环节，并逐渐与人工智能新技术协同发展成为组织的核心业务之一。数据治理与人工智能的发展存在相辅相成的关系。一方面，数据治理为人工智能奠定了基础。通过数据治理，数据质量得以提升，从而为人工智能的应用提供高质量的合规数据。另一方面，人工智能对数据治理存在诸多优化作用。通过人工智能技术，数据治理工作中的数据模型管理、数据质量管理、数据安全等方面均可在智能化水平方面得到相应的提升。

区块链具有高容灾、不可篡改、动态网络扩展、可扩展权限控制等优势，这些基础优势体现在数据应用与数据安全方面，在数据治理中可实现数据分层及优化网络结构。应用区块链的解决方案，将数据写入和读取等权限规则记录到链上，对数据保管进行分层处理，通过其不可篡改和数据冗余的特性，确保数据不会丢失，并且实现了明确的责任划分，解决了事后责任追溯难的问题，确保在数据治理过程中数据的所有权、使用权、存储权等更有保障，增强了数据治理的可靠性。智能合约数据共享使用提供了治理手段。智能合约是一套以数字形式定义的承诺，承诺包含了合约参与者约定的权利和义务，由计算机系统自动执行。基于智能合约可以有效实现针对数据收集、共享、使用等关键环节的自动化治理。

2. 平台赋能与数据治理

在数字经济时代，数据成为新型的关键生产要素，同时也是驱动产业高质量发展的强劲引擎。目前，各产业逐渐向数字化转型，其中最重要的方式之一是建立数字化平台，创新发展平台经济，在新一代数字科技的支撑和引领下，以数据为关键要素，以释放数据价值为核心。这有助于提高数据的准确性，加快数据洞察，打破数据孤岛，通过统一平台，链接、管理所有分散、异构的数据资产，实现数据治理“无死角”。

3. 行业引领与数据治理

作为核心生产要素，数据的核心科学决策作用日渐凸显，数据治理对于各行各业的社会价值和经济作用来说，具有十分重要的意义。释放数据的价值需要构建科学合理的数据治理体系，各类行业作为数据治理的重要参与者，均在不断尝试培育有序有效的数据市场。众多行业，如医疗、教育等，通过数据的有效供给，培育数据新业态，以应用场景牵引带动数据要素市场的繁荣发展，以领域为主体不断探索数据治理体系，不断进行行业升级，以技术为支撑，以市场为纽带，形成良性的数据资源，为数据治理提供新的机遇。

1.4.4 数据治理的发展趋势

数据正重塑着经济与社会。从生产到生活，从工业到服务业，从产业端到消费端，越来越多地呈现出数据化的态势，数据已经开始重新定义一切。在过去几年，数据治理作为数据的核心管理手段，得到了国家、政府、企业、个人的高度关注。伴随着理论、法律、政策、产业的一系列实质性变化，各方正在将数据治理纳入到政务活动、企业治理、经营



视频讲解



管理等领域，数据治理的理念、法规、方法、工具等也得到了蓬勃发展。随着各行业、各组织对数据治理实践的推进，一些变化与趋势正在逐步显现。

趋势一：数据治理成为数字化转型的核心要素。数字经济作为增长新引擎持续发力，其热度不减，“十四五”规划时期，我国数字经济转向深化应用、规范发展、普惠共享的新阶段，“数字化转型”仍是很多企业近年的重要战略部署之一。同时，以云计算、大数据、人工智能、区块链等为代表的数字技术不断涌现，快速向各领域进行深度融合，加快了数字化转型的进程，数据量和数据价值密度都呈现出爆炸式的增长。海量的数据资源，对数据采集、存储、分析、处理的工具、计算、建模应用等方面的数据能力提出了更高的要求。数据治理成为实现数据、技术、流程和组织的职能协同、同台优化和互动创新的主要依托环节。数据治理使数据成为数字化转型的关键驱动要素，赋能国家战略与企业运营的创新发展，深入挖掘数据资产价值。

趋势二：数据治理依托人工智能新技术走向智能化。数据治理和人工智能作为近几年的两大浪潮，二者之间一直存在相辅相成的关系。一方面，数据治理为人工智能奠定了基础。通过数据治理，可以提升数据质量、增强数据合规性，从而为人工智能的应用提供高质量的合规数据。另一方面，人工智能对数据治理存在诸多优化作用。人工智能有助于实现概念模型与计算机模型的完美融合，从而优化数据模型管理；有助于实现对非结构化数据的采集和关键信息的提取，并帮助维护、整合碎片数据；有助于定义转换规则，提取数据质量评估维度，通过监督学习、深度学习来实现对数据清洗和数据质量的效果评估，最大化地实现数据质量的动态提升；有助于推进数据分级分类，促进数据安全保障体系的完善，进一步保障数据安全。

趋势三：数据治理从成本中心向价值中心演进。传统的数据治理往往聚焦于政府或者企业的内部数据能力建设。但在目前数字化转型的大背景下，数据要素的激活、数据价值的发挥、数据服务的建立与开放逐渐成为国家、企业在进行数据治理时的关注重点，数据治理的定位逐渐向价值中心演进，更注重效能。这与数据治理的直接目标是一致的，即挖掘和释放隐藏在数据中的巨大价值。我国绝大多数行业领域和政府部门、企事业单位及其他社会组织的数据在整体数据资源体系中占据着巨大的体量，其数据质量水平也是数据资源体系中最好的部分。这些数据在全面、准确地反映客观事物的性质和状态方面更加值得信赖和依靠，且更易于查找和被获取，更便于被利用。与传统文献形态的信息资源相比，数字化的数据形态更易于被功能强大的信息系统、信息技术高速和有效处理，从而发挥出巨大的作用。而数据治理是数据开发、利用的保障，更是数据开发、利用功效的放大器。

趋势四：数据将进一步开放与共享。在数字经济时代，数据已成为基础性战略资源，作为新型生产要素之一，数据资产化已是必然趋势，而数据的开放共享则是深入挖掘数据资产价值的基础。从2015年9月国务院发布《促进大数据发展行动纲要》至今，政府数据的开放共享正不断推进，各方面资源进行了有效整合，综合治理能力大幅提升。2021年《数据安全法》第五章专章规定，实现“政务数据安全与开放”，在基本形成跨部门数据资源共享共用格局后，由政府主导打通政府部门、企事业单位间的数据壁垒，建立数据共享开放平台，进一步推动实现政府公共数据的普遍开放。2022年，随着全社会的数据存储、数据挖掘、数据使用、数据参与意识逐渐觉醒，数据价值化的条件将进一步成熟，

数据的所有权、使用权、增值权以及数据红利的释放权、分配权有望在新的一年里确定更加清晰的边界，数据要素价值将得到更有效的释放。此外，在数据不断标准化、共享化的同时，行业的数据标准建设进程也将进一步加快，无论是政府、行业数据，还是企业内部数据，都将遵循一个相互认可的数据标准、处理规程。这也将进一步推动国家、企业建立相关相应的数据治理标准、路径与方法。

趋势五：数据安全仍是贯穿数据治理各环节的核心重点。随着数字经济时代的到来，国家、企业和个人对网络的依赖程度不断加深，数据安全成为国家安全和经济社会稳定运行的基础。另外，人们的日常生活工作都离不开网络，个人和法人等主体的身份、财产和活动等都将以数据形态呈现。因此，海量、多元和非结构化已经成了数据发展的新常态，这为数据治理带来了许多安全隐患。数据治理要做到安全先行，因此，数据安全治理仍将是各组织进行数据治理的核心重点之一。数据安全正步入法治化和战略性轨道。截至2021年，我国的数据安全监管框架已经基本成型：《网络安全法》《数据安全法》《个人信息保护法》三法为数据安全护航；此外，在银行、通信、工业、能源等领域也已经有一系列条例规章从实践角度推动产业内数据安全治理体系的落地。数据安全与隐私保护至关重要，目前国家与各行业都在通过政策、法规等完善数据安全的保护路径，逐步建立起多维度、多层面完整的数据安全保护体系，进而更好地推进数据的利用朝着健康的方向发展。

1.5 思考题

1. 请谈谈你对数据价值延伸的看法。
2. 试分析建立数据治理道德准则的必要性和可行性。
3. 请概括总结数据治理发展的现状和趋势。