

第 1 章

绪 论

1.1 数据的相关概念

1.1.1 数据

“数据”一词最早出现在拉丁语中,其最初的含义为“给予的事物”。之后,随着科学技术的发展和人类认知的演进,“数据”一词逐渐进入其他语种,并被应用到更多领域。工业革命以来,人类科学技术进入了高速发展阶段,数据的内涵也不断丰富,人们对数据的理解也更加深刻。近几十年来,互联网、云计算、大数据、人工智能(artificial intelligence, AI)、区块链等新一代信息技术(information technology, IT)不断涌现,并快速渗透到经济社会的各个领域,数据规模急剧扩大,数据形态更加复杂,数据应用更加广泛,数据价值更加凸显,数据治理也面临更加严峻的挑战。

关于数据的内涵理解和概念界定,不同的机构、学者和专业技术人员等在不同时期,从不同视角进行了探究,一些代表性的定义见表 1-1。

表 1-1 关于数据的不同定义

机 构	定 义	时 间
联合国统计委员会(United Nations Statistical Commission, UNSC)、联合国欧洲经济委员会(United Nations Economic Commission for Europe, UNECE)	数据是信息的物理表现形式,这一表现形式适用于人工或自动化手段交流、理解或处理	2000 年
国际空间数据系统咨询委员会(Consultative Committee for Space Data Systems, CCSDS)	数据是以适合于交流、解释或加工的形式化方式进行的可重新解释的信息表示,例如,序列、数值表、页面中的字符、录音等都是数据	2012 年
国际数据管理协会(Data Management International, DAMA)	数据是以文本、数字、图形、图像、声音和视频等格式对事实进行表现的形式,是信息的原始材料	2017 年
美国《开放的、公开的、电子化的及必要的政府数据法案》(The Open, Public, Electronic, and Necessary Government Data Act)	数据是以任何形式或介质记录下来的信息,开放政府数据时特别指明数据需要满足机器可读的条件	2019 年
《深圳经济特区数据条例》《上海市数据条例》等	数据是指任何以电子或者其他方式对信息的记录	2021 年

上述不同时期、不同主体关于数据的概念界定虽然有所差异,但也有一些共同之处,从这些定义中可以归纳出,数据的定义有三个关键要素:数据是一种对信息的记录形式;数据的来源和范围非常广;数据具有一定的价值。

基于不同的准则,对数据有不同的划分。

(1) 根据数据存储媒介不同,可以将数据分为传统物理空间的纸质媒介数据和网络空间的电子媒介数据。纸质图书、档案、文件、报告等记录的都是纸质媒介数据,而光盘、硬盘、闪存盘等存储的电子化数据以及存在于数据中心或云中心的海量数据等都是电子媒介数据。

(2) 基于数据是否表现为数值,可以将数据分为数值型数据和非数值型数据。例如,一个人身高 170 厘米、体重 65 千克,一块土地面积 650 平方米,空气湿度 60%等都是数值型数据;而一段文字、一幅图画、一条语音、一段视频等都是非数值型数据。

(3) 从数据是否连续来看,可以将数据分为连续数据和非连续数据。例如,声音的大小或温度的变化都是连续数据,而符号、图像、文字等都是非连续数据。

(4) 根据数据的表现形式和存储方式,可以将数据分为结构化数据、半结构化数据和非结构化数据。结构化数据通常是指可以存储在关系型数据库中,可以用二维表结构来逻辑实现的行数据,如银行卡号、邮政编码等;非结构化数据则是不使用数据库二维逻辑结构来表现的数据,如文本、图像以及各类文档和多媒体数据等;半结构化数据是介于结构化数据和非结构化数据之间的数据,如超文本标记语言(hyper text markup language,HTML)文档就是一种半结构化数据。

随着互联网、云计算、物联网、区块链、人工智能等新一代信息技术的快速发展,网络空间的电子化数据规模呈现指数型增长,数据形态更加复杂,数据应用更加丰富,互联网时代的大数据通常是指网络空间的电子化数据,数据资产管理的对象也主要是指电子化数据。

网络空间的电子化数据具备以下一些特有的属性。

(1) 物理属性,即数据在存储介质中以二进制串的形式存在。数据的物理存在占据了存储介质的物理空间,是数据真实存在的表现,并且是可度量、可处理的。数据的物理存在使数据可以直接被用于制作副本、进行传输,也可以通过特殊的方法直接从物理存在勘探数据、破解数据。

(2) 存在属性,即数据以人类可感知的形式存在。在网络空间中,物理存在的数据可以通过输入、输出设备以某种形式展现出来,被人所感知、所认识,只有通过输入、输出设备感知到的数据才能被认为是存在的数据。

(3) 信息属性,即数据是否有含义和含义是什么。通常,数据经过解释之后会具有含义,数据的含义就是其所承载和展示的信息;当然,并不是所有数据都具有明确的含义,例如一个随机数字或一串乱码都是数据,但没有明确的含义;也有一些数据,其含义会随着时空变化而演变,一些当前没有明确含义的数据,经过一定时期之后就具有了含义,一些表面没有明确含义的数据,经过处理或转换就具有一定含义。

(4) 时间属性,即数据是否附加时间标记。自然界中的事物通常都具有时间属性,而数据可以没有时间属性,如果数据要被用于表示自然界一个随时间变化的事物,对应于自然界的时间概念,那么需要给数据加盖时间戳。例如“2022 年 5 月 10 日合肥市的最高气温是 28℃”,这里的“28℃”就具有了时间戳“2022 年 5 月 10 日”,有了明确的时间属性。此外,

与自然界中的事物会随着时间老化不同,虽然记录数据的载体会老化,但数据不会“老化”,可以通过更换存放数据的载体来保证数据一直被存储下来。

1.1.2 信息

信息是音讯、消息、通信系统传输和处理的对象,泛指人类社会传播的一切内容,是一切通信和控制系统中的一种普遍联系的形式。信息、物质、能量是组成客观世界的三大基本要素。人们通过获得和识别来自社会与自然界的不同信息来区分不同事物,进而认识和改造世界。

扩展阅读 1-1
从长城、烽火与
驿传制度入手,
浅析我国古代的
信息传递方式



在中国古代,“信息”即“消息”。中国是世界上最早建立信息传递组织的国家之一,从中华文明起源开始,人们利用驿站、乐器、烽燧和飞鸽等设施及手段在人与人、部落与部落之间进行信息传递。自古以来,驿传系统就在文报传递、官员接待、运输中介三个方面发挥着重要作用。随着经济社会的快速发展,以及交通工具的不断进步,信息传递需求更加迫切,信息传递质量要求更高,信息传递方式更加灵活。

1928年,拉尔夫·文顿·里昂·哈特利(Ralph Vinton Lyon Hartley)在《贝尔系统技术杂志》(*The Bell System Technical Journal*)上发表的论文《信息传输》(*Transmission of Information*)对消息和信息进行了区分,指出“信息是指有新内容、新知识的消息”,他将信息理解为选择通信符号的方式,并用选择的自由度来衡量这种信息的大小。1948年,克劳德·艾尔伍德·香农(Claude Elwood Shannon)在《贝尔系统技术杂志》上发表的论文《通信的数学理论》(*A Mathematical Theory of Communication*)指出,“信息是用来消除随机不确定性的东西”,也就是说,信息是用来降低随机不确定性的东西。他利用概率论阐明了通信中一系列基本理论问题,给出了计算信源、信息量和信道容量的方法与一般公式,得到了一组表示信息传递重要关系的编码定理。1950年,控制论的奠基人诺伯特·维纳(Norbert Wiener)认为,正如“熵”是无组织(无序)程度的度量一样,消息集合中包含的信息就是有组织(有序)程度的度量,因此,可以将消息所包含的信息解释为负熵。他在《人有人的用处:控制论与社会》(*The Human Use of Human beings: Cybernetics and Society*)一书中指出,“信息就是我们在适应外部世界,并把这种适应反作用于外部世界的过程中,同外部世界进行交换的内容的名称”。1975年,意大利学者吉乌塞佩·朗格(Giuseppe Longo)在其出版的《信息论:新的趋势与未决问题》(*Information Theory: New Trends and Open Problems*)一书中对信息做了进一步的界定,认为“信息是反映事物的形成、关系和差别的东西,它包含在事物的差异之中,而不是在事物本身”。1996年,中国学者钟义信在《信息科学原理》中对信息的概念进行了详尽的阐述,认为在信息概念的诸多层次中最重要层次有两个,分别为没有任何约束条件的本体论层次和受主体约束的认识论层次,在本体论层次上,信息可被定义为“事物运动的状态以及它的状态改变的方式”;在认识论层次上,信息可被定义为“主体所感知或表述的事物存在的方式和运动状态”,这意味着在认识论层次上,没有主体就没有信息。

信息本身所承载的内涵随着经济社会的快速发展和科学技术的进步而日渐丰富,人们对“信息”的认识和理解也逐渐深入。特别是在计算机和互联网技术发明以后,各行各业产

生了更大规模、更多类型的数据,数据所表达和蕴含的内容通常被认为是重要的信息。区别于其他事物,信息具有一些自身的特性,包括普遍性、时效性、相对性、与物质不可分割性、可传递和干扰性、可加工性和可共享性等。

1. 普遍性

一切事物都有其运动的状态和方式,而信息就是指事物的运动状态和方式。因此,信息随处可见,具有普遍存在性。

2. 时效性

客观事物一般都是运动变化的,信息是事物运动的状态和方式,也在不断变化。因此,一般情况下,信息的存在有一定的时效性。

3. 相对性

客观上信息是无限的,但对于特定认知主体来说,实际能够得到的信息总是有限的。因此,人们感知、获取和利用的信息具有相对性。

4. 与物质不可分割性

信息本身是虚拟的、不可触及的,它必须记录或存储在特定的媒介或载体上。因此,信息不能脱离物质单独存在,通常都具有与物质不可分割性。

5. 可传递和干扰性

信息能够通过多种方式进行传递,信息的价值也往往需要通过传递来实现。一个完整的信息传递过程通常包括信源(信息发送方)、信宿(信息接收方)、信道(信息媒介)和信息四个要素。信息在传递过程中可能会受到各种因素的干扰,产生噪声,或者损失、失真等。

6. 可加工性

信息是可以被加工的,信息加工就是把信息从一种形式变换成另一种形式,为实现信息的价值和基于信息的决策目标,信息可以被分析或综合,被压缩或扩充。由于人为操作或主观因素的介入,信息加工过程往往是不可逆的。

7. 可共享性

与物质和能量的一个显著差异是,相同的信息可以无差别地被不同主体共同使用,信息在共享过程中并不会遭受损失,反而有可能增加新的信息,信息的价值也会通过共享被放大。信息的可传递性也为信息共享搭建了桥梁,使信息的交换、转让和使用可以持续进行下去,从这个角度来看,信息是人类共同的财富。

1.1.3 知识

人类对“知识”的认知与对“数据”和“信息”的认知类似,都随着经济社会的发展和科学

技术的进步经历了一个不断发展的过程。我国悠久的历史文化长河中,积累了相当丰富的关于“知识”概念的探索。《辞源》中就对“知识”给出了两种解释:一是“相识见知的人”;二是“指人对事物的认识”,第二种解释与现代汉语中的含义比较接近。1980年版的《辞海》中对“知识”定义是“人们在社会实践中积累起来的经验”,并指出“从本质上说,知识属于认识的范畴”。《现代汉语词典》(第7版)中将“知识”定义为“人们在社会实践中所获得的认识 and 经验的总和”。

知识管理的奠基人之一卡尔·爱立克·斯威比(Karl-Erik Sveiby)在1997年出版的《新组织财富》(*The New Organizational Wealth*)中提出,知识就是“行动的能力”。在此基础上,1998年,托马斯·H.达文波特(Thomas H. Davenport)等人在《工作知识》(*Working Knowledge*)中指出,知识是“可以辅助我们做出决策或采取行动的有很高价值的一种信息形态”,“知识是一种包含了结构化的经验、价值观、关联信息以及专家的见解等要素的动态的混合物。知识起源于认识者的思想,并对认识者的思想起作用。在组织内,知识不仅存在于文档和数据库中,而且嵌入在组织的日常工作过程、实践和规范中”。通过上述不同阶段、不同视角对知识的定义,可以看出知识具有以下基本特性。

1. 隐性性

知识与其主体密不可分,因为知识总是属于某个人或某个组织的,不同主体掌握的知识是有所差异的,知识之间的交流有时也是不充分的。

2. 实体性

无论是对个人还是对一个组织,知识都是一种实际的资产。知识的生产、集聚和收获过程都需要付出劳动,而且知识往往具有产权,即知识产权。因为知识的实体性,便产生了知识的识别、组织、收集和测度等一系列问题。

3. 共享性

与信息类似,知识也是可以共享的,而且不会随着共享而导致其价值和数量的减少,甚至可能在共享中产生新的知识。知识的共享过程需要转让者和受让者共同参与。

4. 主观性

知识的价值取决于使用者对知识的主观认识、需求和理解,不同使用者对同一知识的价值判断是不同的。

5. 价值性

与数据和信息不同的是,知识更接近行动,因此,知识也更具有价值。根据知识所导致的决策和行动可以评价知识的价值。能够产生更高的生产率、使人们作出更明智决策的知识具有更高价值。

6. 复杂性

知识的组成、结构和变化造成了知识的复杂性。因此,在运用知识时要全面考虑影响知

识的复杂性的各个因素,避免因为对知识复杂性的考察不全面而得出错误的结论。

7. 变化性

知识随着人类社会的不断发展而永恒变化。人类文明起源时,知识就成为人类社会中无处不在、无时不有、不断变化的一种要素。

1.1.4 智能

长期以来,国内外许多哲学家、脑科学家、心理学家以及工程技术专家等一直在努力研究和探索什么是“智能”。

在中国古代,一般认为“智”和“能”是两个相对独立的概念,“智能”是“智力”和“能力”的总称。例如,《荀子·正名篇》中有“所以知之在人者谓之知,知有所合谓之智。所以能之在人者谓之能,能有所合谓之能”。其中,“智”指进行认识活动的某些心理特点,“能”则指进行实际活动的某些心理特点。也有一些古代典籍把“智”“能”二者结合起来看作一个整体。例如,《吕氏春秋·审分览》中有“不知乘物,而自怙恃,夺其智能,多其教诏,而好自以……此亡国之风也”,《论衡·实知篇》中有“故智能之士,不学不成,不问不知”,“人才有高下,知物由学,学之乃知,不问不识”。这里把“人才”和“智能之士”相提并论,认为人才就是具有一定智能水平的人,其实质就在于把“智”与“能”结合起来作为考察人的标志。

近现代以来,国内外学者从不同视角认识“智能”,对其概念内涵的理解进一步深入。他们将知识和智力总结为智能,其中,知识是智能行为的基础,而智力是获取并应用知识求解问题的能力。智能具有以下特征:一是具有感知能力,即能够感知外界,从而获取知识;二是具有记忆和思维能力,记忆用于存储思维所产生的知识,而思维则用于对信息的处理,是获取知识以及运用知识求解问题的根本途径;三是具有学习能力,即通过与环境的作用不断学习;四是具有行为能力,行为能力就是信息的输出,对外界的变化作出反应。

1956年夏,约翰·麦卡锡(John McCarthy)和马文·李·明斯基(Marvin Lee Minsky)等科学家在美国达特茅斯学院(Dartmouth College)就“如何用机器模拟人的智能”这一问题开会研讨,会议首次提出“人工智能”这一概念,这也标志着人工智能学科的诞生。经过60多年的发展,人工智能在算法、算力(计算能力)和算料(数据)“三算”方面取得了许多突破,虽然人工智能从“能用”到“好用”还有很多瓶颈难题,但不可否认“三算”方面的突破有力推动了人工智能的技术进步和行业应用。

互联网时代,数据是人工智能的底层基础设施,数据的有效管理和应用也驱动许多人工智能新理论、新方法和新技术的产生。随着新一代信息技术的快速发展及其在各行各业的加速渗透,各行各业的智能化进程加快,智能化水平不断提升,一些“人工智能”替代或超越了“人类智能”。在智能交通、智慧医疗、智能家居、智能工厂、智能建筑等应用场景下,“人工智能”技术的运用使各类新产品、新服务、新模式不断涌现,人类各种差异化、个性化、灵活化的需求能够更好地被满足。

1.2 数据资产的相关概念

1.2.1 资产

在人类历史长河中,资产的形式不断演变和丰富,从原始社会的石头、木头,到各类植物、动物,再到各类矿产资源和金银珠宝等有形资产,以及商标、品牌、股票、债券、专利、创意等无形资产,每一种资产的形成与发展,都给人类带来了巨大的商业价值和财富机会,造就了各类新颖的商业模式和成功的企业精英。资产价值的发挥离不开科学、有效的资产管理,基于资源本身的价值创造更多价值。

资产是财务会计的一个基本要素,是一个重要的会计学概念。经济学家保罗·萨缪尔森(Paul Samuelson)等人认为“资产是实物财产或具有经济价值的非实物性的权利,如工程、设备、土地、专利、版权,以及货币、债券等金融工具”。不同机构给出了不同的资产定义,总结如表 1-2 所示。现有对资产的定义可以归纳为“资产是由各类会计主体[企事业单位、政府机构、非政府组织(non-governmental organizations, NGO)等]过去的交易或事项形成的、由会计主体拥有或者控制的、预期会给会计主体带来经济利益或产生服务潜力的经济资源”。

表 1-2 关于资产的不同定义

机 构	定 义	时 间
美国注册会计师协会(American Institute of Certified Public Accountants, AICPA)	按会计规则或会计原则进行结账而被结转 为借方余额所代表的某些事物	1953 年
(美国)财务会计准则委员会(Financial Accounting Standards Board, FASB)	某一特定主体由于过去的交易或事项所获 得或控制的可预期的未来经济利益	1985 年
国际会计准则理事会(International Accounting Standards Board, IASB)	企业因过去的交易或事项而控制的资源,这 种资源可以给企业带来未来的经济利益	1989 年
英国会计准则委员会[Accounting Standards Board(UK), ASB]	由过去的交易或事项形成的、由特定主体控 制的、对未来经济利益的权利或其他使用权	1999 年
中华人民共和国财政部	企业过去的交易或者事项形成的、由企业拥 有或者控制的、预期会给企业带来经济利益 的资源	2006 年

个人和组织拥有的资产通常可以分为有形资产与无形资产两类。

1. 有形资产

有形资产是指个人或组织拥有的具有实物形态的资产,包括固定资产和流动资产。如存货、对外投资、货币资产、应收账款等具有实体物质产品形态的资产,包括生产有形资产和非生产有形资产。

2. 无形资产

无形资产是指个人或组织拥有、没有实物形态、可辨认的非货币性资产。无形资产包括社会无形资产和自然无形资产,其中,社会无形资产包括专利权、非专利技术、商标权、

著作权、特许权、土地使用权等；自然无形资产包括不具有实体物质形态的天然气等自然资源等。

资产管理就是围绕各类资产的配置、使用和处置等开展的计划、组织、协调、控制等一系列管理活动，如资产的核算规整、资产的价值评估、资产的维修维护、资产的保值增值、资产的流通交易等。因为有形资产和无形资产本身存在较大差别，所以有形资产和无形资产的管理过程、特点和内容存在较大差异。数据资产是一类特殊的无形资产，数据资产管理与一般无形资产管理相比虽然有一些共性，但数据资产的特性决定了其管理方法、内容、目标等都有很多独特性。

1.2.2 信息资产

随着信息技术的发展和信息化步伐的加快，经济社会各领域的信息化程度不断提高，特别是管理信息系统的普及和应用，各领域产生的数据不断增加，信息资源成为信息化的必然产物。信息资产是个人或组织所拥有或者控制的，能够为其带来未来经济利益的信息资源，其本质在于信息作为一种经济资源参与企业的经济活动，降低和消除了企业经济活动中的风险，能够为个人或组织的管理控制和科学决策提供合理依据，并预期带来经济利益。

对于一个企业来说，其广义的信息资产大致可以分为六类，如表 1-3 所示。从概念内涵边界来看，数据资产是信息资产的一种形式，一些信息资产管理的方法也适用于数据资产管理。但随着互联网时代的到来，与早期的“数据”概念相比，互联网时代数据的形态、规模、结构、价值都发生了深刻变化，数据资产的概念也远远超出了传统信息资产中数据资产的含义，具有更加广泛的内涵和外延。传统的信息资产管理已难以适应互联网时代的数据资产管理，数据资产管理亟须新的理论、方法和技术支撑。

表 1-3 信息资产的分类

分 类	一 般 描 述
软件资产	企业信息基础设施和设备上部署、安装与使用的各种软件，主要用于存储、处理和管理各类信息，包括系统软件、应用软件、工具软件等，例如操作系统软件、网络安全软件、数据库软件、办公自动化软件等
实物资产	支撑企业业务的各类信息设备或硬件基础设施，用于部署软件、存放数据或支持信息管理，包括存储设备、网络设备、计算设备、显示输出设备等
人员资产	对数据资产、软件资产和实物资产进行操作、运维和管理的各类人员，包括企业管理人员、业务操作人员、技术支持人员、开发维护人员等
服务资产	对其他信息资产的运行和管理发挥支撑作用的各类服务，包括技术支持服务、运行维护服务、远程协助服务、安全保护服务等
其他资产	除了上述几类信息资产，能够为业务提供支持的其他无形资产，如标准规范、客户资源、战略规划等
数据资产	以物理或电子的方式记录的数据，如文件资料、电子数据等。文件资料包括公文、合同、操作单以及一些日常数据和外来流入文件等。电子数据包括制度文件、管理办法、技术报告、数据库数据等

1.2.3 数据资源

在信息化的初期,各领域收集到的数据量相对较小,数据形态比较简单,人们对数据价值的认识也相对有限,因此还没有形成“数据资源”的概念。随着信息化不断深入,个人、企业、组织和全社会集聚的数据规模越来越大,数据形态更加复杂,人们对数据的认识也更加全面、深刻,逐渐形成了数据和信息是一种重要资源的认知。基于各行业、各领域的信息化、数字化的特点和需求,集聚形成了面向不同行业、领域的

扩展阅读 1-2
从“资源”到“资产”
用数据资产管理解锁数据
新价值



的数据资源。

1. 个人数据资源

个人数据资源主要是指个体在工作、学习、生活、娱乐等方面的行为所产生并集聚的数据。例如,个人在日常工作和学习中的笔记等文档,在生活中拍摄的照片、视频等。特别是在互联网环境下,自媒体高度发达,个人数据资源的时空范畴进一步扩大,个人在即时通信软件、社交媒体、购物软件、短视频平台等创建个人账号,并发布、浏览、评论的大量数据等,都是广义的个人数据资源。

2. 政府数据资源

政府和公共部门是较早引入信息技术的领域之一。在信息化背景下,政府自身运转和公共服务效率不断提高,在这个过程中积聚了大量数据资源。除了政府政务活动所形成的政务数据资源,政府数据资源还包括由财政资金支持而形成的公共数据资源,如各行业统计数据、国家和地区统计年鉴、各类公共服务数据库等。特别是在数字政府背景下,政府数据资源的范畴得到了极大的拓展。

3. 企业数据资源

与政府和公共部门类似,各类企业也不断融合信息技术,以提高自身的管理水平和经营业绩,并变革企业发展战略。企业在生产经营、供应链管理、客户服务等过程中,集聚形成大量内部、外部数据资源,如产品或服务数据、客户数据、营销数据、市场信息等,这些数据资源对支撑企业精准、高效、科学的管理决策至关重要。特别是在电子商务环境下,企业在互联网上开展的各种线上营销、获得的在线评论等成为企业数据资源新的重要来源。

4. 医疗数据资源

医疗数据资源包括医院在诊疗、治疗,以及预防保健等过程中,关于医患主体、医疗服务过程和医疗服务效果等的各类数据记录。除了医生开具的诊断报告和处方、患者的病历档案、各类医学检查影像和化验结果等传统医疗数据外,医疗数据资源还包括基因组学数据、医疗保险数据、药物研发数据、医学研究数据等。特别是在智慧医疗环境下,各类人体传感器、医疗机器人和智能诊疗设备等能够采集到更大规模、更加复杂的医疗大数据。

5. 交通数据资源

在交通行业,各类交通工具、交通基础设施和交通管理部门等的运行和管理活动产生的数据是交通数据资源的直接来源。与交通相关领域的业务系统数据是交通数据资源的间接来源。例如,交通导航系统数据、交通能源保障系统数据和社交媒体上关于交通的交流讨论数据等。特别是在智慧交通环境下,人一车一路一环境交互协同更加复杂,交通大数据形态、规模和价值等得到极大拓展。

1.2.4 数据资源的特征

当前,信息技术已经渗透到我们经济社会的各个领域,融入生产生活的各个方面,数据资源增长速度更快、存在形态更加多样、应用价值更加多元,人们对数据资源的特征和价值的认知更加全面、深入,有人认为数据资源的价值可以和煤炭、石油、天然气等自然资源相提并论,甚至在某些方面超过自然资源。可以认为,数据资源是能够反映物质世界和精神世界运动状态及其变化的一类信息资源,它具有复杂性、决策有用性、高速增长性、价值稀疏性、可重复开采性和功能多样性,一般具有多种潜在价值。

1. 复杂性

与其他信息资源相比,数据的形式和特征更加复杂,数据资源的数量规模、主要来源、形态结构、状态变化和开发方式等都是造成其复杂性的原因。

2. 决策有用性

虽然数据资源直接利用效率有限,但可以通过对其进一步分析和挖掘发现其中隐含的知识,在实际应用中提供决策支持,这是其他资源难以具备的能力,也是数据资源价值在决策有用性方面的体现。

3. 高速增长性

与自然资源不同,数据资源的总存量不仅不会随着人类不断地开采而逐渐减少,反而还会迅速增加,而且是呈指数型增加,甚至可能出现爆发性态势。例如在互联网上,通过搜索引擎、社交媒体和电子商务等方式,每秒会产生大量的数据。因此,在利用数据资源支持管理决策时,面临的一个主要挑战就是时效性问题。

4. 价值稀疏性

数据资源规模巨大,在给资源利用带来机遇的同时,也伴随着一定挑战。其中一个主要挑战就是数据资源价值的低密度问题,即数据资源规模虽大,但其中蕴藏的有用价值却是很稀疏的,这为数据资源的开发和利用增加了难度。

5. 可重复开采性

在这一特性上,数据资源与自然资源也有很大不同。由于开发利用过程的不可重复性,自然资源的存量会随着资源的不断开采而减少。但数据资源是可以被重复开发利用的,并

且对于给定的数据资源,凡是具备该资源使用权的人或组织都可以进行开采和挖掘,而且这些数据资源在被使用之后,仍可以由其他人或组织继续开发和利用。

6. 功能多样性

特定的自然资源在人类生产和生活中的功用有一定的局限性,而对于数据资源,其功能往往随着资源开发目的和方式的不同而发生改变,例如数据资源可以用于风险预警、趋势预测、异常识别、模式挖掘、营销和服务模式创新等。需要注意的是,即使数据资源具有上述诸多积极功能,但也要注意避免因数据资源的不当或非法利用而给个人、企业和组织以及国家和社会带来威胁。

1.2.5 数据要素化

传统的经济学中,将资本和劳动力作为最基本的生产要素,其中,资本涵盖了土地、能源资源等“物”的要素,而劳动则是指“人”的要素,“人”不断地开发和利用“物”,并将它们转化为商品和服务,这就是推动经济增长的本质。在信息化初期,数据总量是非常有限的,数据还不具备成为生产要素的条件。

然而,进入互联网时代,万物互联、智能感知、泛在连接使得全球范围内的数据量飞速增长,数字经济蓬勃发展,数据的要素特征开始显现,数据具备了成为生产要素的基础条件,数据要素在经济社会中的关键作用也逐渐加强。培育发展数据要素市场已成为要素市场改革的重要组成部分和发展高质量数字经济的基础。2019年10月31日,中国共产党第十九届中央委员会第四次全体会议通过的《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》中明确指出,“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”,首次明确了数据是参与分配的生产要素之一,这标志着由数据这一全新生产要素引发的数字经济进入新的发展阶段,将重塑生产和消费、供给和需求的关系,变革个体行为方式、组织运作模式和社会治理体系。2020年3月30日,《中共中央 国务院关于构建更加完善的要素市场化配置体制机制的意见》中又进一步指出:要加快数据要素市场的培育,包括推进政府数据开放共享,提升社会数据资源价值,加强数据资源整合和安全保护等,发展数据要素市场是未来数字经济健康、快速、持续发展的基础。2022年3月25日,《中共中央 国务院关于加快建设全国统一大市场的意见》提出:“加快培育数据要素市场,建立健全数据安全、权利保护、跨境传输管理、交易流通、开放共享、安全认证等基础制度和标准规范,深入开展数据资源调查,推动数据资源开发利用。”

数据作为一类生产要素,与劳动力、土地、资本等传统生产要素相比,主要有以下三个方面的独特性。

1. 非稀缺性

由于资源往往具有稀缺性,因此经济学通过寻求以最小投入获得最大产出的方式处理社会上的有限资源与人的无限欲望之间的关系。自然资源和劳动人口都是有限的资源。数据资源则与它们有所不同,虽然数据也不是人们可以随心所欲获取和使用的资源,但数据资源总体是非稀缺的。当前新一代信息技术仍处在快速发展和应用阶段,因此,数据量是急剧

增长的,而大规模数据资源的开发利用会受到存储和计算等物理设施的约束。数据资源的非稀缺性还体现在数据的可重复开发利用方面,即数据资源的规模不会因为开发利用而减小,反而往往会增加。

2. 非均质性

劳动力、土地、资本等传统生产要素具有一定的均质性。例如,劳动力通常在一定时空范围内是相对均质的,差异性并不明显;单位价值的资本、单位面积的土地之间往往也没有明显差别。而对数据要素来说,两个规模相同但来源不同的数据价值可能完全不同,甚至没有可比性。因为容量规模只是数据的一个度量维度,数据资源的价值更多地体现在数据的产生途径、存在形态和处理方式等上。

3. 非排他性

传统的劳动力、土地、资本等生产要素通常都具有排他性,即这些要素被某些主体使用后就不能再被其他主体使用。例如,一个企业拥有一个煤矿的开采权,其他企业就不能再进行开采;一个企业拥有一块土地的使用权,其他企业就不能再使用。而数据要素具有明显的非排他性特征,在产权明晰和使用合规等前提下,数据可以被多个主体同时开发和利用。随着与数据安全和数据管理相关的法律法规不断完善,数据的开发和使用时也将面临更多约束。

1.3 数据资产化

1.3.1 数据资产

数据符合资产的定义,具备资产的基本特性。在经济社会运行过程中,各类主体能够通过采集、处理、加工、购买等方式拥有并控制数据资源,再通过转让数据、挖掘数据或共享数据等服务获得收益。因此,数据资源是能够给各类主体带来收益、可以被各类主体拥有或控制的资源,其可以由各类主体过去的交易或者事项形成,这就使数据资源具备成为资产的基本条件。而如果一类数据资源的成本和价值能够被可靠地计量,便符合资产的要求,也就可以作为数据资产。数据资产可以定义为:各类主体合法拥有或控制的一类以物理或电子方式记录的数据资源,这类资源能够进行计量、交易或流通,并能持续发挥作用,直接或间接地为各类主体带来价值。

数据资产是有别于传统实物资产和其他类别无形资产的一种资产形态,从定义中可以看出数据资产具有以下特点:一是资源性,即数据资产是一类数据资源,可以用物理方式或电子方式记录,主要是文本、图像、语音、视频、网页、数据库、传感信号等结构化、半结构化和非结构化数据;二是归属性,即数据资产是个人或组织合法拥有的数据资源,既包括组织内部数据,也包括第三方的外部数据,但都有明确的所有权;三是价值性,即数据资产是能够为其所有主体带来价值的数据资源,价值既包括有形的经济效益,也包括无形的品牌价值、声誉价值和其他社会价值等;四是运营性,即数据资产是能够计量的资产类型,并且能够进行交易和流通,通过运营产生价值。

1.3.2 数据资产化的内涵

数据资产管理包括数据资源化和数据资产化两个阶段。其中,数据资源化是将原始数据转变为数据资源,使数据具备一定的潜在价值,进而构建全面有效、切实可行的数据资产管理体系,并提高数据质量,保障数据安全的过程,是数据资产化的必要前提;

扩展阅读 1-3
数据资产化的发
展与挑战



而数据资产化则是在数据资源化的基础上,进一步丰富数据资产应用场景,建立数据资产生态,持续运营数据资产,实现数据资产的业务价值、经济价值和社会价值的过程。

数据资源化以数据治理为工作重点,以提升数据质量、保障数据安全为目标,确保数据的准确性、一致性、时效性和完整性。数据资源化包括数据模型管理、数据标准管理、数据质量管理、主数据管理、数据安全治理、元数据管理(meta data management)、数据开发管理等管理过程。数据资产化通过将数据资源转变为数据资产,使数据资源的潜在价值得以充分释放。数据资产化以扩大数据资产的应用范围、显性化数据资产的成本与效益为工作重点,并使数据供给端与数据消费端之间形成良性反馈闭环。数据资产化主要包括数据资产流通、数据资产运营、数据价值评估等管理过程,是数据创造价值的过程,也是发挥数据要素价值、培育数据要素市场的必由之路。

并非所有的数据都是资产,只有那些在合法合规的前提下,各类主体(包括企业、政府和各类组织)有权利、有能力控制和使用的,且有潜力为这些主体带来价值(包括商业价值和社会价值)的数据,才可以被定义为数据资产。数据资产可以通过交易和事项这两类方式形成。

1. 通过交易的方式形成数据资产

目前,各级地方政府在探索通过交易形成数据资产,积极建立数据交易示范机构,例如,上海数据交易所、贵阳大数据交易所等数据交易机构。通过交易的方式形成数据资产的优势在于市场化手段和机制能够有效支撑数据的确权和流通,通过市场化定价能够明晰数据资产的价值。数据所有权交易的环节包括身份认证、服务协议、数据资产管理登记与发布审核、数据资产评估、交易磋商与合同签署、交付与验收、交易结算。

2. 通过事项的方式形成数据资产

在信息技术的支撑下,各类主体通常是通过信息化事项,积累和形成数据资源的。在形成数据资源的基础上,再通过数据确权、数据定价等途径,就可以逐步完成数据资产化,最终加工成数据产品并在市场上流通。

数据资产与传统资产所具有的资产特性有所不同,因而数据在其资产化过程中,面临着数据资产形态、数据权属以及数据资产评估与定价机制等方面的挑战。在数据资产化过程中,要从法律法规、管理机制、财务计量、技术支撑等多个方面进行探索,从而形成科学合理、完整规范的数据资产化方法,不能使用现有的会计体系直接处理数据资产,也不能简单地将传统的资产管理手段运用到数据资产领域。

案例 1-1 上海数据交易所

2014年,上海市出台《2014年度上海市政府数据资源向社会开放工作计划》,通过分析多个领域中数据的需求度和成熟度,结合社会需求,详细制定了当年的总体思路、重点工作任务和工作要求,并向社会公开。此外,上海市公共数据开放平台还支持用户对数据集进行评分,并对平台的数据查询、数据获取、数据展示等功能提出建议。

上海数据交易所于2021年11月25日在上海市浦东新区成立,是为贯彻落实《中共中央国务院关于支持浦东新区高水平改革开放打造社会主义现代化建设引领区的意见》中的重要任务,由上海市人民政府的相关部门和机构推动组建,面向全球开展大数据综合交易,旨在推动数据要素流通、释放数字红利、促进数字经济发展。上海数据交易所采用公司制架构,围绕打造全球数据要素配置重要枢纽节点的目标,构建“1+4+4”体系:紧扣建设国家级数据交易所“一个定位”;突出准公共服务、全数字化交易、全链生态构建、制度规则创新“四个功能”;体现规范确权、统一登记、集中清算、灵活交付“四个特征”。上海数据交易所面向数据流通交易提供高效便捷、合规安全的数据交易服务,同时引导多元主体加大数据供给,培育发展“数商”新业态。

2022年1月1日起施行的《上海市数据条例》提出公共数据授权运营机制,参照公共资源特许经营的模式,由上海市人民政府办公厅采用竞争方式确定被授权运营主体,运营主体可在授权期限和范围内以市场化方式运营公共数据,提供数据产品、数据服务并获得收益。获得公共数据授权运营的机构由上海市大数据中心实施日常监管,由上海市人民政府办公厅组织制定公共数据授权运营管理办法,明确授权标准、条件和具体程序要求,建立授权运营评价和退出机制。

资料来源:上海数据交易所.关于我们[EB/OL].[2022-11-15]. <https://www.chinadep.com/about>.

案例 1-2 贵阳大数据交易所

作为中国首个国家大数据综合试验区,贵州省已率先建成中国第一个省级政府数据集聚、共享、开放的系统平台,以大数据与实体经济深度融合赋能传统产业转型,设立了全球第一个大数据交易所。贵州省还制定了《贵州省大数据发展应用促进条例》,这是中国首部大数据地方法规,将大数据产业发展纳入法制轨道。2017年,贵州出台《贵州省政府数据资产管理登记暂行办法》,以加强全省政务服务实施机构数据资产的管理,真实反映全省政府数据资产状况,保证政府数据资产的信息完整、全面、准确,这也标志着贵州省成为全国首个出台政府数据资产管理登记办法的省份。

贵阳大数据交易所是中国第一家以“大数据”命名的交易所,于2015年4月14日完成首批大数据交易。2022年5月27日,贵阳大数据交易所正式对外发布数据交易规则。该规则采用自动计价模式,若数据买方应约价不低于卖方挂牌价,则按照交易所自动撮合成交,成交价为买方应约价格;对于不能自动成交的应约,卖方可选择能接受的应约完成交易,成交价为买方应约价;系统还对数据设定了拆分原则,以应对数据买方不一定需要全部数据样本的情况,这时系统会自动报价,而后自动撮合成交。贵阳大数据交易所基于底层数据,通过对数据进行清洗、分析、建模,再使结果可视化,实现了原始数据“可用不可见”、数据产品“可控可计量”、流通行为“可信可追溯”,解决了数据交易在隐私保护和数据所有权方面的问题。

资料来源:贵阳大数据交易所.贵阳大数据交易所[EB/OL].[2022-11-15]. <https://www.gzdex.com.cn/>.

1.4 数据资产管理活动职能

1.4.1 数据资产管理概述

数据管理的概念由来已久,其内涵也不断丰富,最早可以追溯到对传统非电子化的各类载体记录的符号、文字、图案等信息的管理。20 世纪中叶以来,随着计算机和互联网相关技术的发明与快速发展,计算机系统采集和存储了越来越多的电子化数据,亟须有效的数据管理技术,各类数据库技术的发明就是支撑数据管理的有效方式。随着信息技术的不断发展及其在各行各业的深入应用,各类面向行业和领域业务的信息系统诞生,如办公自动化(office automation,OA)系统、企业资源规划(enterprise resource planning,ERP)系统、客户关系管理(customer relationship management,CRM)系统、人力资源管理(human resource management,HRM)系统以及各类管理信息系统(management information system,MIS)等,数据的形式和内容更加丰富,数据量不断增长,数据应用需求更加多元,这一时期数据管理更多聚焦业务领域和决策目标。随着云计算、物联网、人工智能、边缘计算、第五代移动通信技术(5th generation mobile communication technology,5G)等新一代信息技术的快速出现和广泛渗透,数据呈现出规模性(volume)大、多样性(variety)、价值性(value)、高速性(velocity)等新特征,人们对数据价值的认知更加深刻,数据管理面临的新挑战更大,数据的资源性和资产性特征愈加明显,数据资产管理的理念、内涵、目标和价值更加清晰。数据资产管理的发展历程如图 1-1 所示。

扩展阅读 1-4
大数据环境下,
如何做好企业数
据资产的管理



图 1-1 数据资产管理的发展历程

数据资产管理是在数据资产化背景下,在数据管理基础上的进一步拓展和升级。数据资产管理与传统数据管理的主要区别体现在三个方面。

1. 管理视角不同

数据管理是对数据资源的收集、整理、组织、存储、传输和加工过程进行信息挖掘,并将这些信息作为行动和决策的依据。数据资产管理强调数据是一种资产,并对这种资产的价值、成本、收益等方面展开全生命周期的管理。

2. 管理职能不同

传统数据管理包含数据治理、数据架构管理、数据开发、数据操作管理、数据安全

数据质量管理、参考数据和主数据管理以及元数据管理等职能。而在数据资产管理中,数据标准管理也被纳入数据资产管理职能,针对数据的应用场景需求和平台建设情况,还增加了数据价值管理等职能。

3. 管理要求不同

在数据资源向数据资产转变、数据资源管理向数据资产管理转变的背景下,相应的管理组织架构和制度规范等也在发生转变。与传统数据管理相比,数据资产管理要以更完善的组织架构、更专业的管理团队和更细致的管理制度来确保其规范性、安全性和有效性。

在互联网时代,对个人或组织来说,并非所有的数据资源都是数据资产,只有那些能够为个体或组织带来直接或间接价值的的数据资源才是数据资产。将数据资源转化成数据资产实现价值创造,以及实现数据资产的保值增值,都离不开有效的数据资产管理。数据资产管理就是为实现特定的数据资产增值目标,支持数据驱动的管理决策,对数据资产及其价值创造过程进行的计划、组织、领导、协调、控制等一系列管理活动。数据资产管理的内涵和外延不是一成不变的,而是随着人们对数据以及数据资产认知的不断深化而发展演化的。数据资产管理的技术、范围和目标等也随着数据规模、数据理念、数据挖掘技术、数字经济模式等的发展而演变。数据资产管理通过系统化、全过程的管理方式实现数据的可得、可用和好用等目标,并实现数据资产价值的最大化,以解决组织在数据价值挖掘过程中所面临的诸多问题。

1.4.2 数据资产管理基本架构

随着人们对数据认知的不断深入,以及数据规模的持续增长和数据形态的日益复杂,数据管理的内容也在不断发展和丰富。关于数据管理的理论框架,国际数据治理研究所(Data Governance Institute,DGI)于2004年提出了数据治理框架(Data Governance Framework),如图1-2所示;DAMA于2009年发布了数据管理知识体系(Data Management Body of

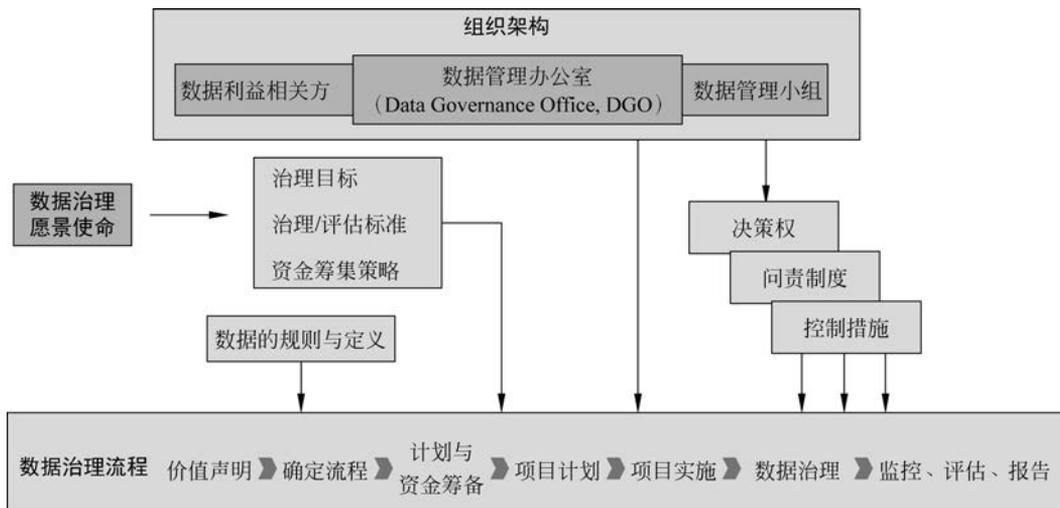


图 1-2 DGI 提出的数据治理框架

Knowledge, DMBOK), 将数据管理划分为数据治理、数据架构管理、数据开发、数据操作管理、数据安全、参考数据和主数据管理、数据仓库和商务智能管理、文档和内容管理、元数据管理和数据质量管理 10 个领域; 2017 年, DAMA 对数据管理模型进行了更新, DMBOK 2.0 将其扩展为 11 个管理职能, 分别是数据架构、数据模型与设计、数据存储与操作、数据安全、数据集成与互操作、文件和内容、参考数据和主数据、数据仓库和商务智能、元数据(meta data)、数据质量等。此外, 高德纳(Gartner)公司、国际商业机器公司(International Business Machines, IBM)等企业也相继提出了数据管理能力评价模型。

2018 年 3 月, 《数据管理能力成熟度评估模型》(Data Management Capability Maturity Assessment Model, DCMM)(GB/T 36073—2018)发布, 成为首个数据管理领域的国家标准。DCMM 是一个综合框架, 从标准规范、管理方法论、评估模型等方面对数据管理进行多维评估。DCMM 的核心评价维度及能力域如图 1-3 所示。

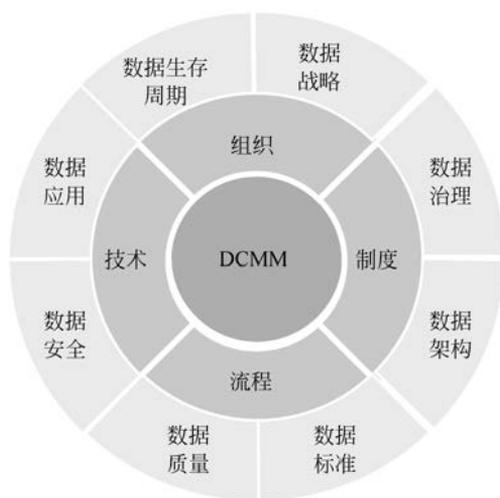


图 1-3 DCMM 的核心评价维度及能力域

DCMM 将组织内部数据能力划分为 8 个重要组成部分, 描述了每个组成部分的定义、功能、目标和标准。其中, 8 个核心能力域对应的 28 个能力项如表 1-4 所示, DCMM 等级划分及主要特征如图 1-4 所示。

表 1-4 DCMM 8 个核心能力域及 28 个能力项

能力域	能力项	能力域	能力项
数据战略	数据战略规划	数据应用	数据分析
	数据战略实施		数据开放共享
	数据战略评估		数据服务
数据治理	数据治理组织	数据安全	数据安全策略
	数据制度建设		数据安全策略
	数据治理沟通		数据安全审计
数据架构	数据模型	数据质量	数据质量需求
	数据分布		数据质量检查
	数据集成与共享		数据质量分析
	元数据管理		数据质量提升

续表

能力域	能力项	能力域	能力项
数据标准	业务术语	数据生存周期	数据需求
	参考数据和主数据		数据设计和开发
	数据元		数据运维
	指标数据		数据退役

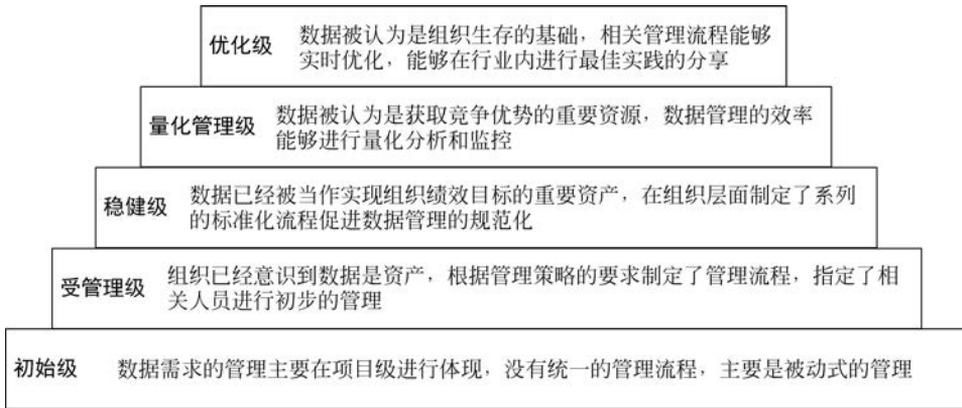


图 1-4 DCMM 等级划分及主要特征

数据资产管理内容丰富,包括数据资产化框架设计、数据资产归类、数据资产确权、数据资产评估、数据资产定价、数据资产运营、数据资产保护等。其中,数据资产评估包括质量评估、价值评估、风险评估等。从数据资源化和数据资产化两个阶段的角度,可以进一步划分数据资产管理的内容,理解数据资产管理,其过程和架构如图 1-5 所示。

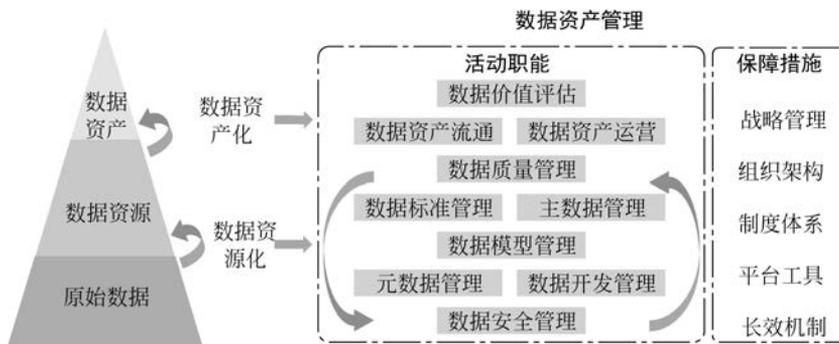


图 1-5 数据资产管理的过程和架构

1.4.3 数据标准管理

数据标准是一种规范性约束,用来保障数据的内外部使用和交换的一致性与准确性。数据标准应用包括数据模型设计与开发、数据质量稽核等。数据标准管理是结合制度约束、过程管控、技术工具等手段制定和发布由数据利益相关方确认的数据标准,以达到推动数据

标准化、提升数据质量的目标。

数据标准管理的关键活动包括：①数据标准管理计划，即明确数据标准管理负责人与参与者，采集并调研数据标准需求和现状，构建组织级数据标准分类框架，制定并发布数据标准管理规划与实施路线；②数据标准管理执行，即基于数据标准分类框架定义数据标准，依据数据资产管理认责体系进行数据标准评审并发布，依托平台工具在数据模型设计与开发、数据质量稽核等环节应用数据标准；③数据标准管理检查，即及时检查数据标准的适用性、全面性，依托平台工具检查并记录数据标准应用程度；④数据标准管理改进，即通过制定数据标准维护与优化的路线图，遵循数据标准管理工作的组织结构与策略流程，各参与方配合进行数据标准维护与管理过程优化。

1.4.4 数据模型管理

数据模型用于抽象现实世界的的数据特征，描述一组数据的概念和定义。数据模型管理是指在信息系统设计阶段就参考逻辑模型，使用标准化用语、单词等数据要素设计数据模型，并在信息系统建设和运行维护过程中，严格按照数据模型管理制度，审核及管理新建和存量的数据模型。数据模型根据其应用范畴分为组织级数据模型和系统应用级数据模型。组织级数据模型包括主题域模型、概念模型和逻辑模型三类，采用组织架构指导建立组织级数据模型，并采用一体化建模的方法，是提升数据模型业务指导性和质量的有效方式。系统应用级数据模型包括逻辑数据模型和物理数据模型两类。

数据模型管理的关键活动包括：①数据模型计划，即确认数据模型管理的相关利益方，采集、定义和分析组织级数据模型需求，遵循数据模型标准与要求设计数据模型；②数据模型执行，即参考逻辑数据模型开发物理数据模型，保留开发过程记录，根据数据模型评审准则与测试结果，经数据模型管理的参与方评审无异议后发布并上线该模型；③数据模型检查，即确定数据模型检查标准，定期开展数据模型检查，保留数据模型检查结果，建立数据模型检查基线，确保数据模型与组织级业务架构、数据架构、信息技术架构的一致性；④数据模型改进，即根据数据模型检查结果，召集数据模型管理的相关利益方，明确数据模型优化方案，持续改进数据模型设计方法、架构、开发技术、管理流程、维护机制等。

1.4.5 数据质量管理

数据质量指在特定的业务环境下，数据满足业务运行和管理决策的程度，是保证数据应用效果的基础。数据质量管理是为了衡量、提高和确保数据质量的规划、实施与控制等而运用相关技术手段进行的一系列活动。衡量数据质量的指标体系包括完整性、规范性、一致性、准确性、唯一性、及时性等。

数据质量管理的关键活动包括：①数据质量管理计划，即确定数据质量管理相关负责人，明确数据质量的内部需求与外部要求，参考数据标准体系定义数据质量规则库，构建数据质量评价指标体系，制定数据质量管理策略和管理计划；②数据质量管理执行，即依托平台工具管理数据质量内外部要求、规则库、评价指标体系等，确定数据质量管理的业务、项

目、数据范畴,开展数据质量稽核和数据质量差异化管理;③数据质量管理检查/分析,即记录数据质量稽核结果,分析问题数据产生原因,确定数据质量责任人,出具质量评估报告和整改建议,持续监测全流程数据质量,监控数据质量管理操作程序和绩效,确定与评估数据质量服务水平;④数据质量管理改进,即建立数据质量管理知识库,完善数据质量管理流程,提升数据质量管理效率,确定数据质量服务水平,持续优化数据质量管理策略。

数据质量管理要遵循“源头治理、闭环管理”的原则。其中,源头治理主要指采用“数据开发管理一体化”理念,在新建业务或 IT 系统过程中,明确数据标准或质量规则,并与数据生产方和数据使用方确认,常见于对数据时效性要求不高或核心业务增量数据等场景。闭环管理主要指形成覆盖数据质量需求、问题发现、问题检查、问题整改的良性闭环,如图 1-6 所示,对数据采集、流转、加工、使用全流程进行质量校验管控,根据业务部门数据质量需求持续优化质量管理方案、调整质量规则库,构建数据质量和管理过程的度量指标体系,不断改进数据质量管理策略。

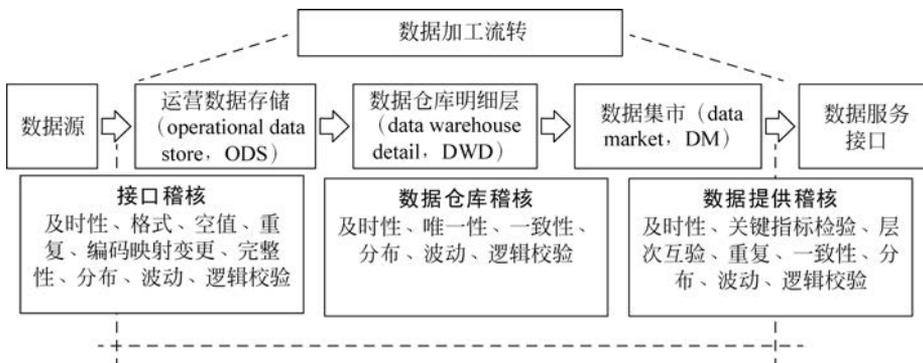


图 1-6 数据全流程质量校验管控

1.4.6 主数据管理

主数据是组织中需要跨系统、跨部门共享的核心业务实体数据。主数据管理是对主数据标准和内容进行管理,实现主数据跨系统的一致、共享使用。

主数据管理的关键活动包括:①主数据管理计划,即定义主数据的数据模型、数据标准、数据质量、数据安全等要求和规则,并明确以上各方面与组织全面数据资产管理的关系,依据数据模型明确主数据的业务范围、唯一来源系统与识别原则;②主数据管理执行,即通过平台工具实现核心系统与主数据存储库的数据同步共享;③主数据管理检查,即检查主数据的质量并记录主数据检查的问题,以保证主数据的一致性、唯一性;④主数据管理改进,即总结主数据管理问题,制订主数据管理提升方案,以提升主数据质量及管理效率。

主数据具有价值高、稳定性强、数量少但影响范围广等特点,将主数据管理作为数据资产管理的切入点,可以覆盖数据资产管理全流程,因此主数据也有“黄金数据”之称。与主数据所描述的对象不同,参考数据是指用于对其他数据进行分类的数据。参考数据管理是对定义的数据值域进行管理,包括采用标准化术语、代码值和其他唯一标识符对每个取值的业

务定义,对数据值域列表内部和不同列表之间的业务关系进行控制,以及实现相关参考数据的一致、共享使用。

1.4.7 元数据管理

元数据是关于数据或数据元素的数据,以及关于数据拥有权、存取路径、访问权和数据易变性的数据。元数据管理是关于元数据的构建、存储、整合与控制等一整套流程的集合。元数据管理是数据资产管理的重要基础,是为获得高质量、整合的元数据而进行的规划、实施与控制行为。

元数据管理的关键活动包括:①元数据管理计划,即明确元数据管理的参与方,采集元数据管理需求,确定元数据类型、范围、属性,设计元数据架构,并制定元数据规范;②元数据管理执行,即采集和存储元数据,可视化数据血缘,通过非结构化数据建模、自动维护数据资产目录等应用元数据;③元数据管理检查,即元数据质量检查与治理,实施元数据治理执行过程规范性检查与技术运维,保留元数据检查结果并建立元数据检查基线;④元数据管理改进,即根据元数据检查结果,召集相关利益方并明确元数据优化方案,制订改进计划以持续改进元数据管理的方法、架构、技术与应用等内容。

元数据贯穿于数据资产管理的全流程,是支撑数据资源化和数据资产化的核心。首先,元数据从业务视角和管理视角出发,通过定义业务元数据和管理元数据,增强业务人员和管理人员对数据的理解与认识。其次,技术元数据通过自动从数据仓库、大数据平台、提取-转换-加载(extract-transform-load,ETL)中解析存储和流转过程,追踪和记录数据血缘关系,及时发现数据模型变更的影响,有效识别变更的潜在风险。最后,元数据可作为自动化维护数据资产目录、数据服务目录的有效工具。图数据库可以有效提升元数据血缘关系的存储和分析效率,具有良好的展示实体关联关系的特性,且技术相对成熟,已逐步应用至元数据血缘分析中。

1.4.8 数据开发管理

数据开发是指将原始数据加工为数据资产的各类处理过程。数据开发管理是指通过建立开发管理规范与管理机制,面向数据、程序、任务等处理对象,对开发过程和质量进行监控与管控,使数据资产管理的开发逻辑清晰化、开发过程标准化,增强开发任务的复用性,提升开发的效率。依托统一的数据开发平台,可以从技术侧和管理侧提升数据开发管理效率。

数据开发管理的关键活动包括:①数据开发管理计划,即制定数据集成、开发和运行维护规范;②数据开发管理执行,即设计集数据集成、程序开发、程序测试、任务调度、任务运维等功能于一体的数据开发工具,根据数据集成规范进行数据的逻辑集成或物理集成,根据数据使用方的需求进行数据开发;③数据开发管理检查,即监控数据处理任务的运行情况并及时处理各类异常;④数据开发管理改进,即定期进行数据集成、开发、运维工作复盘,并以此为基础对相关规范持续更新迭代。

案例 1-3 中国工商银行数据开发管理

中国工商银行从 2002 年开始建设数据集市,当时主要使用 Oracle 类单机版的关系型数据库。随着数据量不断增加,2014 年,中国工商银行正式基于 Hadoop 技术建设了大数据平台,构建企业级数据湖及数据仓库。2017 年,随着 AI 技术的兴起,其又开始建设机器学习平台,2020 年开始建设数据中台和高时效类场景。

中国工商银行搭建了大数据开发工作站和研发与测试管理系统,对数据开发过程进行效率管控。在大数据开发工作站中创建生产工作区,与常规生产运行资源、数据资源等解耦隔离,构建“端到端”的数据服务流水线。在现有 Hive、MPPDB 等批量加工的基础上,进一步满足流式数据加工、联机数据访问服务的开发场景,将编程语言由 SQL(结构查询语言)向 Spark、Python 等扩展。研发与测试管理系统建立了数据开发需求管理指标,包括各需求项的平均周期、开发前置时间、开发节奏等。此外,采用“统计过程控制”(statistical process control,SPC)的理念,使用统计方法对开发过程与任务进行实时质量监控。相较于 2020 年第三季度,2021 年第三季度中国工商银行数据需求的平均研发周期大幅缩短,数据需求响应效率提升 60%左右。

资料来源:CCSA TC601 大数据技术标准推进委员会,中国信息通信研究院云计算与大数据研究所.数据资产管理实践白皮书(5.0 版)[R]. 2021.

1.4.9 数据安全治理

数据安全是指通过采取必要措施,确保数据被有效保护、合法利用,并且能够保障数据持续安全。数据安全治理是指在组织数据安全战略的指导下,多个部门协作确保数据安全状态的一系列活动,包括建立组织数据安全治理团队,制定数据安全相关制度规范,构建数据安全技术体系,建设数据安全人才梯队等。

数据安全治理的关键活动包括:①数据安全治理计划,即理解组织内外部数据安全需求与监管要求,制定涵盖数据安全工作的基本原则、数据安全治理规则和程序、内外部协调机制等方面的数据安全治理制度体系,并且明确处理规则、合规审计制度、跨境传输安全评估体系等个人信息保护管理制度,定义并发布数据分类分级标准规范;②数据安全治理执行,即依托平台工具识别敏感数据,应用数据安全分类分级标准规范,根据数据的敏感级别部署相应的数据安全防控系统或工具,例如,权限管控、数据脱敏、数据防泄露、安全审计等;③数据安全治理检查,即监控数据在采集、存储、传输、加工、使用等环节的安全、隐私及合规状况等,并进行组织内外部数据安全审计;④数据安全治理改进,即总结数据安全治理问题与风险,评估数据安全治理相关标准规范的适用性、有效性,持续优化数据安全治理过程。

数据安全分类分级是数据安全治理的基础性、关键性工作。《中华人民共和国数据安全法》(以下简称《数据安全法》)从国家层面明确了数据安全分类分级对于建立数据安全制度的重要性,其中第 21 条明确提出“国家建立数据分类分级保护制度”。此外,2019 年 5 月正式发布的《信息安全技术 网络安全等级保护基本要求》(GB/T 22239—2019)中提出网络运营单位“应根据资产的重要程度对资产进行标识管理,根据资产的价值选择相应的管理措施”;“应对信息分类与标识方法作出规定,并对信息的使用、传输和存储等进行规范化管

理”。2020年3月30日,《中共中央 国务院关于构建更加完善的要素市场化配置体制机制的意见》中明确提出“推动完善适用于大数据环境下的数据分类分级安全保护制度,加强对政务数据、企业商业秘密和个人数据的保护”;金融、证券、工业、电信等行业都出台了相应的数据分类分级指南,以数据资产分类为基础,结合敏感数据分级规则,形成数据资产安全分类分级标准。2016年9月28日,贵州省质量技术监督局发布《政府数据 数据分类分级指南》(DB 52/T 1123—2016),定义了政府数据的分类分级原则和方法。该指南按照科学性、稳定性、实用性、扩展性的分类原则,将政务数据划分为主题、业务、服务三种类型,同时,根据数据敏感程度将数据划分为非敏感数据、涉及用户隐私数据、涉及国家秘密数据三类,形成不同数据资产类型下的数据等级管控要求。图 1-7 为数据安全分类分级流程与结果。

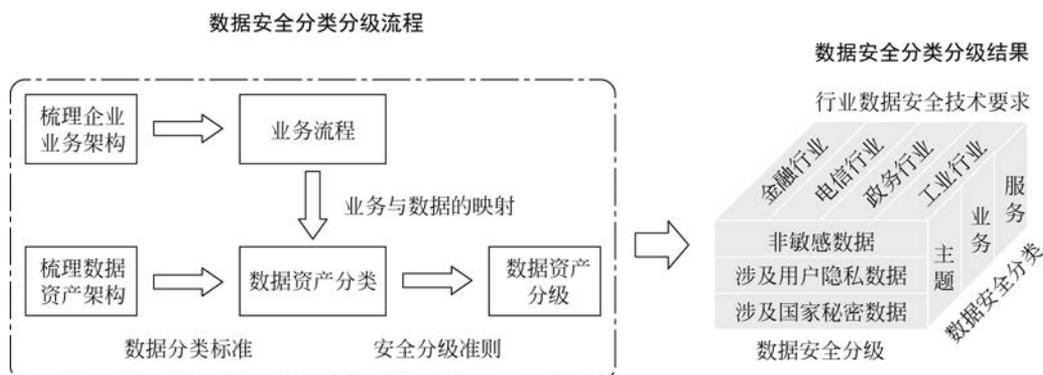


图 1-7 数据安全分类分级流程与结果

此外,个人信息保护是数据安全的重要内容。《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》)将自然人姓名、出生日期、身份证件号码、生物识别信息等全面纳入保护范围,为组织确定了个人信息保护范畴与要求。数据安全要系统地识别业务涉及的个人信息处理活动,充分掌握个人信息收集、存储、流通等活动,并将个人信息保护作为数据安全标准规范的建立依据。数据资产流通要在确保个人信息安全的前提下,引入安全多方计算(secure multi-party computation, SMPC)、联邦学习(federated learning, FL)等隐私计算技术。

案例 1-4 厦门市大数据安全开放平台

大数据安全开放平台是能够为社会公众提供统一业务办理网站的大数据安全开放网站系统,是解决数据主体数据被使用的授权、鉴权、存证的统一第三方授权系统,是解决安全开放过程规范化问题、可追溯管理的大数据安全开放管控系统,还是解决安全计算问题的大数据安全开放安全算力系统。厦门市大数据安全开放平台由厦门市信息中心建设管理,该平台为社会公众、企业获取可公开的政府数据资源提供了统一通道,平台第一批开放了 39 个部门 800 余万条数据记录、789 个数据集、327 个服务接口,涵盖信用服务、交通运输、市场监管、生态环境、地理空间、生活服务等 22 个领域主题。

厦门市大数据安全开放平台不仅是全国首个大数据安全开放平台,也是全国首个采用数据“安全屋”技术开展政务大数据安全开放应用的平台,使海量的数据可以安全合规开放。数据“安全屋”技术通过沙箱技术、堡垒机、审核流程等保证数据安全性,从网络、数据、业务

多层次建立数据安全保障机制,实现了数据所有权与使用权分离,做到数据“可用不可见”。数据“安全屋”技术还支持多样且丰富的数据源,以创造更高价值。用户可以在线完成数据申请、数据授权、数据分析、操作监控和审计、结果使用等操作。

在大数据安全开放平台上,依据公共数据资源敏感程度,可以将数据开放分为普遍开放和授权开放。为进一步提高数据隐私安全保障,更好地保护数据主体的权益,厦门市大数据安全开放平台还搭建了基于区块链技术的统一第三方授权平台,实现了基于区块链的全流程审计、基于计算沙箱的实时操作监控和历史回放、基于结果使用规范的访问限制,全方位保证数据的使用安全,确保数据授权过程安全可信。厦门市利用该平台使政务大数据安全开放,打破了信息的壁垒,让数据提供方更有安全感,让数据需求方更有获得感,为政务大数据开发利用提供了保障。基于该平台已衍生出了政务决策、普惠金融、商业选址、交通出行、医疗卫生等多个场景化应用。

资料来源:厦门市大数据安全开放平台. 平台简介[EB/OL]. [2022-11-20]. https://data.xm.gov.cn/opendata/other/#/about_platform.

1.4.10 数据资产流通

考虑交换数据属性和交换主体的不同,可以将数据资产流通划分为数据共享、数据开放和数据交易等形式。数据共享是指建立统一的数据共享机制,打通组织各部门间的数据壁垒,加速数据资源在组织内部的流动。数据开放是指组织向社会公众提供易于获取和理解的数据。特别地,对于政府而言,数据开放是指公共数据资源开放,对于企业而言,数据开放是指公开企业运行情况、推动政企数据融合等。数据交易是指交易双方在安全合规的前提下,通过合同约定,开展以数据或其衍生形态为主要标的的交易行为。在数据交易中,并非所有交易都以货币进行结算,在遵循等价交换的前提下,“以物易物”“以数易数”或“以数易物”等方式都可能存在。

对于具备公共属性的数据,在组织体系内部流通属于数据共享,如政府机构之间的数据交换,在组织体系外部流通属于数据开放,如公共数据向社会公众开放。对于具有私有(商品)属性的数据,在组织内部流通属于企业数据共享,如企业部门间的数据交换,在组织外部流通属于数据交易。数据共享和数据开放应充分考虑数据使用者的需求,形成需求清单,明确数据流通的合规要求和潜在安全风险,建立数据安全分类分级标准,利用技术工具对数据资产流通过程开展安全防护。在数据采购方面,确定各业务线的数据采购需求,明确数据采购的必要性,构建可信外部数据分布图谱,并对数据供应商进行可信度评估,集中管理采购数据类别及报价。在数据交易方面,评估当前数据交易环境,制定年度数据服务产品清单,形成基本的数据定价标准、交易流程标准、交易合同模板,使用交易平台或数据流通技术执行交易过程,同时,根据数据交易市场的变化,优化数据交易模式、定价规则、交易平台。

案例 1-5 青岛海洋数据交易平台

为解决海洋数据供需不匹配问题,提升数据共享意识,打通数据交易渠道,满足科研院所、海洋物联网企业等主体的数据需求,推动实现海洋数据多样化应用,青岛市大数据发展管理局支持青岛国实科技集团有限公司建设海洋数据交易平台,打破上下游数据壁垒,带动多源数据融合和深度价值挖掘,培育第三方数据服务商,活跃行业数据交易市场。

海洋数据交易平台面向各类海洋科研机构以及相关企业,开展海洋地质、地形地貌、水文气象、遥感影像等海洋数据交易。平台利用区块链技术精选数据确权,实现数据所有权和使用权分离,为数据确权探索新路径;通过系统监测和专家知识并行确保数据质量,提高数据审核效率,确保数据“可用、好用”;构建海洋数据资产价值评估体系预估数据价值,结合实际交易过程确定交易价格,为数据资产定价提供参考。截至2022年7月,青岛国实科技集团有限公司海洋数据交易平台已有地波雷达实时观测、海洋预报、海洋牧场观测三类数据,数据存储量约100 T,初步实现了数据交易,交易金额达30万元,激发了海洋数据市场活力。

资料来源:中国信息通信研究院.中国数字经济发展报告(2022年)[R].2022.

1.4.11 数据价值评估

数据价值评估是指通过构建价值评估体系,计量数据的经济效益、业务效益、投入成本等活动,是数据资产管理的关键环节。狭义的数据价值是指数据的经济效益,广义的数据价值是在经济效益之外,考虑数据的业务效益、成本计量等因素。从内在价值、成本价值、经济价值、市场价值四个价值维度出发,可以建立数据资产价值评估体系。

1. 内在价值

内在价值是指数据本身所蕴含的潜在价值,通过数据规模、数据质量等指标进行衡量。评估数据资产内在价值是评估数据资产能力的基础,对数据资产其他维度的价值评估具有指导作用。

2. 成本价值

成本价值是指数据获取、加工、维护和管理所需的财务开销。数据资产的成本价值包括获取成本、加工成本、运维成本、管理成本、风险成本等。数据资产成本价值评估可用于优化数据成本管理方案,有效控制数据成本。

3. 经济价值

经济价值是指对数据资产的运用所产生的直接或间接的经济收益,通过货币化方式计量数据资产为企业带来的效益。数据资产的经济价值评估可以指导数据资产管理策略,为组织数据资产管理提供决策支持。

4. 市场价值

市场价值是指在公开市场上售卖数据产品所产生的经济收益,由市场供给决定数据资产价值。随着数据产品需求的增加以及数据交易市场规则的建立,从市场价值维度建立数据资产价值评估体系的可行性与准确性逐步提升。

1.4.12 数据资产运营

通过数据资产运营,对数据服务、数据流通等情况进行持续跟踪和分析,以数据价值评估为参考,从数据使用者的视角出发,全面评价数据应用效果,建立科学的正向反馈和闭环

管理机制,促进数据资产的迭代和完善,适应和满足数据资产的应用和创新需求。

统一的数据服务平台可以屏蔽底层数据的技术细节,在底层数据平台升级或迁移过程中,减小对业务的影响,从而提高数据链路构建和运行的效率,缩短数据使用者获取数据的时间,减少数据在不同角色中传递的信息损耗。因此,数据资产运营可以通过使用统一平台、丰富数据使用形式、构建数据生态等途径提供数据服务,进而提升数据服务效率,提升数据资产运营效果。此外,政府还可以采取公共数据授权运营的方式,由市场主体作为数据运营管理方或数据交易中介,缓解政府部门公共数据运营压力,提升公共数据运营效率。

1.5 数据资产管理的保障措施

1.5.1 战略管理

扩展阅读 1-5
数据资产管理
概要



战略是为指导组织持续稳定发展而进行的长期发展规划和资源配置等一系列行动。在互联网时代下,数据战略已成为组织开展精益数据资产管理的基础,指导数据资产管理工作的长期高效开展。如图 1-8 所示,数据战略管理是指通过对数据战略进行规划、执行和评估,确立数据资产管理的中长期目标和管理活动优先级,明确需要的资源投入总量和资源分配机制,并使数据战略始终契合组织的业务战略。

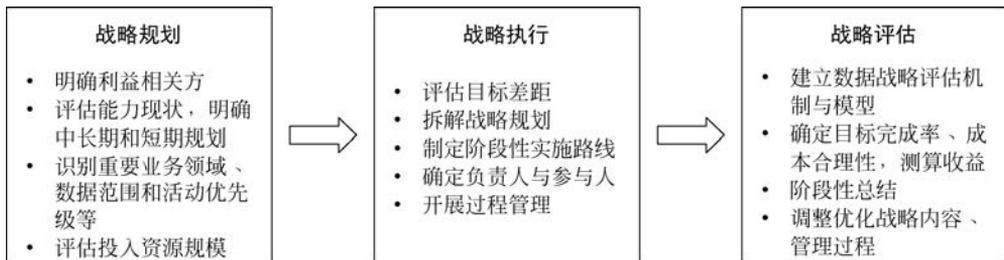


图 1-8 数据战略管理流程与要点

数据战略规划是战略管理的首要环节和基础性工作,是数据资产管理的指导蓝图。要通过评估组织当前的数据资产管理能力,结合组织发展规划和 IT 发展规划,明确数据战略规划的中长期和短期发展目标、管理原则、主要活动,识别重要业务领域、数据范围和活动优先级,确认投入的资源规模。数据战略执行作为战略管理的中间环节,是战略规划落地的有效保障。数据战略执行通过对数据战略规划进行拆解,制定阶段性提升计划与实施路线,明确各项活动参与团队,并根据实际执行情况及时调整短期数据战略规划。数据战略评估是优化组织数据战略管理、提升数据战略指导作用的必要手段。通过从目标完成率、价值收益率、成本合理性等维度评估数据战略内容、管理过程,提升数据战略规划和数据战略执行的有效性。

1.5.2 组织架构

实施组织级统一化、专业化数据资产管理的基础就是建立全方位、跨部门、跨层级的数

据资产管理组织架构,这也是数据资产管理责任落实的保障。一般来说,数据资产管理组织架构包括决策层、组织协调层、数据资产管理层、工作执行层四个层级:①决策层作为数据决策方,由组织首席信息官(chief information officer,CIO)、首席数据官(chief data officer,CDO)担任,负责制定数据资产管理决策、战略和考核机制;②组织协调层由虚拟的数据资产管理委员会承担,负责统筹管理和协调资源,细化数据资产管理的考核指标;③数据资产管理层由数据资产管理办公室承担,作为数据资产管理的主要实体管理部门,负责构建和维护组织级架构,包括业务架构、数据架构、IT架构等,制定数据资产管理制度体系和长效机制,定期开展数据资产管理检查与总结,并向组织协调层和决策层汇报;④工作执行层由业务部门和IT部门共同承担,负责在数据项目中落实数据资产管理工作,与数据资产管理层协同参与各项活动。

在人员方面,数据资产管理对人才的复合型能力提出了更高的要求。首先,数据资产管理相关人员应具备良好的数据架构、数据安全、法律法规相关技能,涉及数据架构师、数据安全人员、数据合规人员等角色。其次,数据资产管理相关人员应具备较高的业务理解能力,涉及数据标准管理员、数据质量管理员、主数据管理员等角色。再次,数据资产管理相关人员需具备基本的市场洞察与运营能力,与数据生态多方、数据使用者建立良好的合作互动机制。最后,由于数据资产在交易市场中的产品设计、定价规则、流通技术需求增多,数据资产管理相关人员需具备一定的金融学、经济学、密码学知识。

集中式管理与联邦式管理是数据资产管理的两种典型组织模式,主要区别在于数据管理专员集中于数据资产管理层或分布于各个业务部门。集中式数据资产管理组织架构如图1-9所示,集中式管理对各业务线数据独立性要求较低、数据相关性要求较高,采用数据

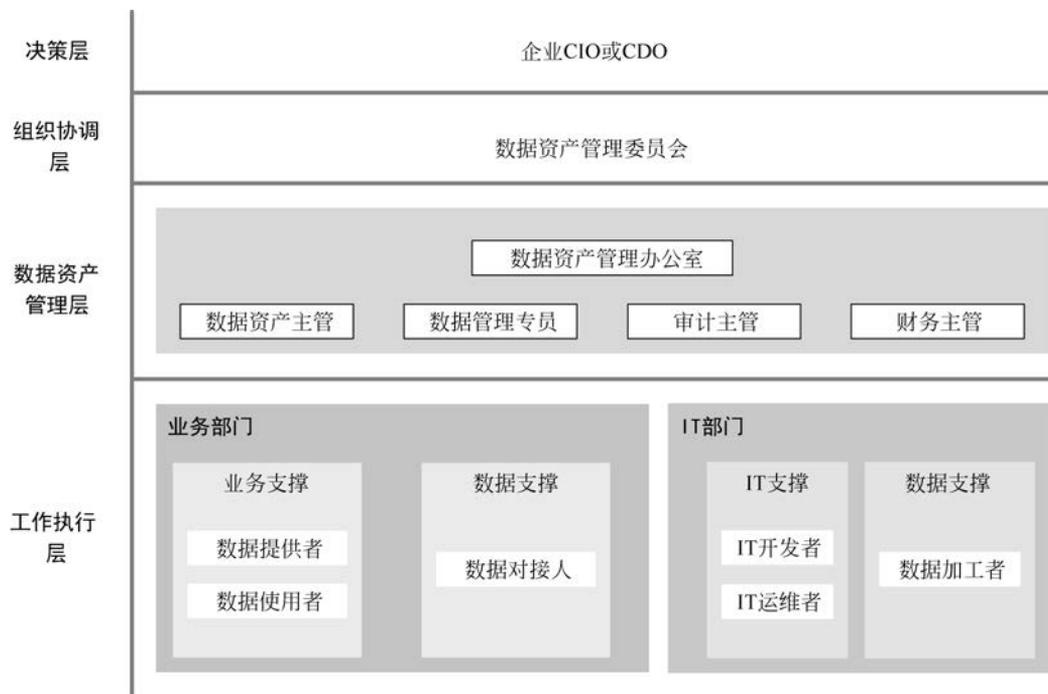


图 1-9 集中式数据资产管理组织架构

仓库、大数据平台等技术,适用于中大型企业。联邦式数据资产管理组织架构如图 1-10 所示,联邦式管理对各业务线数据独立性要求较高、数据相关性要求较低,采用数据经纬(data fabric)技术,适用于中小型企业或集团型企业。如表 1-5 所示,集中式管理与联邦式管理在特点、技术、适用对象等方面存在差异。



图 1-10 联邦式数据资产管理组织架构

表 1-5 集中式管理与联邦式管理比较

组织形式	特点	优势/劣势	适用企业	适用技术
集中式	各业务线独立性较低、数据相关性较高 数字技能在各个业务线分布有所不均	优势: 组织级统一数据资产管理; 全面提升数据资产管理能力 劣势: 与业务结合不足, 敏捷性较低; 投入资源大、见效慢	中大型企业	数据仓库、大数据平台
联邦式	各业务线独立性较高、数据相关性较低 数字技能在各业务线分布相对均衡	优势: 与业务结合, 敏捷性较高; 投入资源小、见效快 劣势: 数据管理团队人员缺口大、技能培养投入大	中小型企业、集团型企业	data fabric

在实践层面,这两种管理和组织模式并没有严格意义上的优劣之分,集中式并不意味着完全的集中管理,联邦式也不意味着完全分散的管理。组织要根据自身数据资产管理基础能力与组织架构决定选用某一种模式,也可以采用集中式与联邦式融合的模式。在理论层面,目前联邦式管理逐渐成为趋势。一方面,由于场景化数据资产应用越发普遍,数据分散管理需求正在增长,从业务端构建数据资产管理团队将有助于理解业务的数据需求,数据直接服务于业务,进而大幅提升数据价值时效性。另一方面,随着数据规模扩大和复杂性的持

续提高,基于 data fabric 的计算和存储架构为数据分散管理提供了技术支持。相较于集中式处理,这种分散处理方式能在一定程度上降低成本,并辅助组织在错综复杂的业务关系网中建立准确、全面、清晰的数据分布图谱。

1.5.3 制度体系

数据资产管理制度体系通常分层次设计,依据管理的颗粒度,制度体系可划分为组织级数据资产管理总体规定、管理办法、实施细则和操作规范四个层次。由于数据产生于业务,数据流转依赖业务,因此,数据认责机制对于统筹协调数据资产管理、加强各方对于数据资产管理的认可具有较强的积极作用。

图 1-11 为数据资产管理制度体系架构,基本内容包括:①数据资产管理总体规定,从数据资产管理决策层和组织协调层视角出发,包含数据战略、角色职责、认责体系等,阐述数据资产管理的目标、组织、责任等;②管理办法,从数据资产管理层视角出发,规定数据资产管理各活动职能的管理目标、管理原则、管理流程、监督考核、评估优化等;③实施细则,从数



图 1-11 数据资产管理制度体系架构

据资产管理层和工作执行层的视角出发,围绕管理办法相关要求,明确各项活动职能执行落实的标准、规范、流程等;④操作规范,从工作执行层的视角出发,依据实施细则,进一步明确各项工作需遵循的工作规程、操作手册或模板类文件等。

案例 1-6 商业银行数据资产管理制度体系

平安银行大数据平台存放银行所有的数据,是业务挖掘的核心资产。为解决数据开发与用户隐私之间的矛盾,平安银行修订了《平安银行数据安全管理办法(2.0版)》,并启动无感数据安全项目,以保障数据安全、提升开发效率。根据基础数据、指标数据、外部数据等数据类别设定数据认责原则,确定不同参与角色在数据生命周期各环节的责任,建立数据认责矩阵,并将数据认责机制与考评激励措施相结合,进一步确保数据资产管理相关方的“责权利”对等。

2019年,交通银行正式成立了数据管理与应用部,负责牵头建立企业级数据治理体系,推动企业级数据中台建设,以数据赋能全行数字化转型。交通银行为响应理论本地化、实践策略化需求,在总体规定、管理办法的基础上,建立了“工作手册”机制,作为数据资产管理工作落地的重要抓手和关键突破口。工作手册覆盖数据资产管理多项活动职能,主要从各项活动实施的需求背景、方案策略、执行过程等方面着手,对自身实践项目进行客观陈述、深入分析,总结问题与经验。工作手册不仅沉淀为内部培训的实战材料,形成团队的共有知识和经验,更重要的是作为典型业务案例,对于长期指导和优化数据资产管理工作具有重要意义。2021年,交通银行制定了《交通银行股份有限公司“十四五”时期(2021—2025年)数据治理规划》,由董事会战略委员会负责对数据治理规划执行情况进行定期监督和评估。

资料来源:CCSA TC601 大数据技术标准推进委员会,中国信息通信研究院云计算与大数据研究所.数据资产管理实践白皮书(5.0版)[R]. 2021.

1.5.4 平台工具

数据资产管理平台工具位于大数据平台上层,为各项数据资产管理活动职能的执行提供技术保障。数据资产管理平台工具在各项数据资产管理活动职能中按照规划—执行—检查—处理(plan-do-check-act,PDCA)流程保障过程流畅、安全。

从管理视角出发,数据资产管理平台工具支持 PDCA 流程:①规划环节,通过角色分配和权限管理落实数据认责体系,支持需求管理,以及对数据资产规模、分布、可信度、安全性等现状进行评估;②执行环节,支持标准规范的新增、修改,以及数据开发、任务编排、任务运维等;③检查环节,支持对数据模型一致性、标准规范应用程度、问题数据处理情况、数据安全响应结果等进行跟踪;④处理环节,支持逐个标记问题并生成改进建议,统计检查数据,形成知识库,量化改进过程,实现闭环管理。

从开发视角出发,数据资产管理平台呈现一体化形式。通过打通数据模型管理、数据标准管理、数据质量管理、数据安全、元数据管理、数据开发相关平台工具,支持数据模型设计与开发遵循标准规范,实现数据质量源头管理,并对数据资产开发全流程进行监控,确保开发过程流畅,提升开发过程的规范性。

1.5.5 长效机制

数据资产管理是一项长期性、战略性、系统化的工作,为保证各项数据资产管理活动有效开展,统筹推动数据资产管理工作顺利进行,不断巩固和加强数据资产管理效能,分工清晰、标准明确、流程规范、奖罚分明的长效机制至关重要。在数据战略管理、组织架构和制度体系的基础上,培训宣贯、绩效考核、激励机制、审计机制、数据文化等长效机制,是数据资产管理活动持续高效运行的重要保障措施,如表 1-6 所示。

表 1-6 数据资产管理长效机制

长效机制	具体措施
培训宣贯	设置培训课程
	内外部经验交流
	优秀案例分享
绩效考核	设计评价指标
	定期组织检查
	检查关联薪酬
激励机制	建立发展路线
	纳入现有晋升机制
	设立相关评比奖项
审计机制	组建设计团队
	审计岗位职责、制度体系、数据活动
数据文化	优化数据服务方式
	提升数据素养
	培养数字技能

1. 培训宣贯是数据资产管理理论落地实践、流程执行运作的基础

通过组织员工参与数据资产管理培训课程、案例分享,进而加强行业内、外部单位先进经验沟通与交流,促进员工提升技术水平。通过建立多层次、多形式、全方位的宣传和培训体系,加强员工对数据资产管理的认识。

2. 绩效考核是确保数据资产管理各项工作落实到位的关键举措

通过建立数据资产管理绩效考核机制,可以检验数据资产管理各个环节的效果。通过日常考核与定期考核相结合、自动考核与人工考核相结合等方式,开展常态化、全面性问题巡检,将问题处理结果与员工薪酬关联,以确保数据认责体系的有效执行。

3. 激励机制是提升组织数据资产管理部门工作积极性,推动数据资产管理良性发展的重要手段

建立员工职业发展通道,设立数据资产管理相关奖项,将数据资产管理纳入现有晋升、薪酬、职位资格等体系范畴。

4. 审计机制是保障数据资产管理按既定规划和规范执行的有效方式

由审计部门、监管部门牵头,数据资产管理部门、技术部门、业务部门参与组建审计团队,引入第三方审计机构,依托相关审计平台,对岗位职责、制度体系、管理活动开展审计。

5. 数据文化是组织开展数据资产管理的核心价值观和最终驱动力

优化数据服务方式,降低数据资产管理参与门槛,开展多类型数据技能培训和比赛,可以加深员工对数据的认识,提升员工对数据的兴趣。

本章小结

不同时期、不同主体对于数据的概念界定有所不同。数据资产的概念已经超出了传统信息资产中数据资产的定义,具有更加广泛的内涵和外延。基于各行业、各领域的数字化特点和需求,目前已形成了面向不同行业、领域的数字资源,这些数字资源具有复杂性、决策有用性、高速增长性、价值稀疏性、可重复开采性和功能多样性等特征。与劳动力、土地、资本等传统生产要素相比,数据要素具有非稀缺性、非均质性和非排他性等方面的独特性。数据资产管理包括数据资源化和数据资产化两个阶段,其中,数据资源化是数据资产化的必要前提,而数据资产化是发挥数据要素价值、培育数据要素市场的必由之路。本章还着重介绍了数据资产管理与传统数据管理的主要区别,数据资产管理的过程和架构,以及数据标准管理、数据模型管理、数据质量管理、主数据管理、元数据管理、数据开发管理、数据安全管理等数据资产管理的主要活动职能。本章总结了战略管理、组织架构、制度体系、平台工具、长效机制等数据资产管理保障措施,以保证各项数据资产管理活动有效开展,支撑数据资产管理实践。

习题

1. 如何理解数据资产与传统信息资产的区别?
2. 结合实例说明数据资源的主要特征。
3. 如何理解数据作为新型生产要素所具备的独特性?
4. 可以从哪些方面衡量数据质量?
5. 简要阐述数据资产管理的各个阶段之间的关系。
6. 参考数据、主数据、元数据三者的关系和区别有哪些?
7. 分别从数据资源化和数据资产化角度阐述数据资产管理的过程与架构。
8. 结合实例分析在数据资产管理中如何保障数据安全。
9. 数据资产价值评估应注意哪些方面?
10. 如何理解数据资产管理的内涵?

即测即练

