

# 从数据的属性看数据资产

朱扬勇<sup>1,2</sup>, 叶雅珍<sup>1,2,3</sup>

1. 复旦大学计算机科学技术学院, 上海 201203; 2. 上海市数据科学重点实验室, 上海 201203;
3. 东华大学计算机科学与技术学院, 上海 201620

## 摘要

在大数据背景下,数据作为数字经济的关键要素已经得到广泛认可,有必要理清信息资产、数字资产和数据资产等概念。综述了信息资产、数字资产、数据资产等概念的发展,分析讨论了数据自身及数据资产的物理属性、存在属性和信息属性,提出将信息资产、数字资产、数据资产等统一为数据资产,并认为数据资产是拥有数据权属、有价值、可计量、可读取的网络空间中的数据资源,数据资产兼有无形资产和有形资产、流动资产和长期资产的特征,是一种新的资产类别。

## 关键词

数据资产;数据资源;大数据;数字经济

中图分类号:TP311

文献标识码:A

doi: 10.11959/j.issn.2096-0271.2018062

## *Defining data assets based on the attributes of data*

ZHU Yangyong<sup>1,2</sup>, YE Yazhen<sup>1,2,3</sup>

1. School of Computer Science, Fudan University, Shanghai 201203, China
2. Shanghai Key Laboratory of Data Science, Shanghai 201203, China
3. School of Computer Science, Donghua University, Shanghai 201620, China

## *Abstract*

In the background of big data now, it has been widely recognized that data is the key factor of digital economy. Therefore, it is necessary to understand the meaning behind of the definitions of information assets, digital assets and data assets. A survey of information assets, digital assets, and data assets was given. The physical attributes, existence attributes and information attributes of data and data assets were discussed. Based on these attributes, information assets, digital assets and data assets were merged into data assets. Data assets were defined as valuable, measurable and accessible data resources in cyberspace owned by an accounting subject. According to the definition and attributes of data assets, data assets have the characteristics of both tangible assets and intangible assets, current assets and long-term assets. Therefore, data assets should be considered as a new category of assets.

## *Key words*

data asset, data resource, big data, digital economy

## 1 引言

2004年,牺牲的美军士兵Justin的父亲John请求雅虎公司告知其儿子在雅虎的账号和密码,以便获取儿子在雅虎账号中留下的文字、照片、E-mail等数据,以寄托对儿子的思念。雅虎公司以隐私协议为由拒绝了该请求,John无奈之下将雅虎公司告上法庭。这个事件引起了公众对个人数据财产、数据遗产的高度关注,可以说是一个历史性事件。

有关网络空间(cyber space)的资产有信息资产(information assets)、数字资产(digital assets)和数据资产(data assets)3个概念。这3个概念名词不同,但内涵类似,主要是由于不同的经济社会发展阶段,形成了不同的名词。这3个概念从名词出现到概念定义,都经历了较长时间,分别如下。

- 计算机创造出来后,使用的技术术语是“信息技术”,信息技术改变了人类生产和生活的方式,“信息资产”一词于1977年由斯图亚特·卡巴克(Stuart Kaback)在介绍一款索引系统时提及,但没有给出具体内涵;1994年,在《霍利报告(The Hawley report)》中首次给出了“信息资产”的定义。

- 数字地球等概念的提出使数字化得到了当时广泛的关注,“数字资产”一词于1996年由Helen Meyer在《维护数字资产技巧》一文中提出,但没有给出具体内涵;2006年阿尔伯特·范·尼凯克(Albert Van Niekerk)给出了“数字资产”的定义。

- 在大数据时代,人们的关注点是数据,数据被作为一种重要战略资源来看待,“数据资产”一词由理查德·彼得斯(Richard Peters)于1974年提出,是指

政府债券、公司债券和实物债券等资产;2009年托尼·费希尔(Tony Fisher)在《数据资产(The data asset)》中指出数据是一种资产。

当前,在大数据背景下,数据是数字经济的关键要素,其作为基础性资源、生产资料已经得到广泛认同,世界主要国家已经在实施大数据战略。从1974年最早出现“数据资产”术语到现在已经40多年了,数据资产进入会计报表仍然问题很多。这说明,未来工作重点应该是如何将信息、数字、数据资源进行资产化。因此,有必要理清信息资产、数字资产和数据资产的异同,定义数据资产的内涵和边界。本文综述了信息资产、数字资产、数据资产等概念的发展,分析了3种资产的概念形成的社会背景和存在的问题,给出了笔者对新的数据资产概念的定义,展望了有关数据资产权属、计量计价方法、会计类别等后续工作。

## 2 相关概念发展状况

本节首先讨论概念的发展状况。按照信息资产、数字资产和数据资产定义出现的时间顺序,分别介绍相关概念。

### 2.1 信息资产

“信息资产”最早由斯图亚特·卡巴克(Stuart Kaback)于1977年介绍一款索引系统时提及,他认为该系统是一种无价的信息资产<sup>[1]</sup>。1981年,福雷斯特·伍迪·霍顿(Forest W. Horton)在《信息资源管理:它从何而来,将何去何从(Information resources management: where did it come from and where is it going)》文章中指出信息资产与其他资产存在重大差异<sup>[2]</sup>。20世

纪90年代,有两份具有广泛影响力的报告:《霍利报告》<sup>[3]</sup>和《信息作为一种资产:无形的金矿》(Information as an asset: the invisible goldmine)<sup>[4]</sup>,这两份报告都将信息视为一种“资产”。1994年,在英国工业联合会(CBI)的支持下,毕马威会计师事务所IMPACT项目发布了《霍利报告》,认为信息是一种重要资源,并将信息资产定义为:“已经或应该被记录的具有价值或潜在价值的信息”,这是信息资产最早的定义。霍利委员会(The Hawley Committee)建议,应该将一个组织中所有的重要信息都识别出来,将其作为企业的资产。信息资产应按其价值和重要性进行识别和分类,以便更好地被管理和利用,确保能够给组织带来最大的商业利益,与管理其他资产(如财产、工厂)一样,董事会应该履行其在信息资产管理的职责<sup>[3, 5-6]</sup>。

1995年,英国路透社的《信息作为一种资产:无形的金矿》一文报道了对500名英国公司高管进行电话采访的调查情况。有近25%的公司表示,信息是其最重要的资产;50%的公司认为,他们的品牌和商标不如他们的信息有价值;超过40%的公司认为,他们没有意识到信息的真正价值,企业希望利用他们自己的信息;但约25%的受访者表示,他们无法将信息资产资本化,因为他们发现很难确定这些资产的价值<sup>[4, 6-7]</sup>。国际标准ISO/IEC 27000:2018(E)中,信息被定义成一种资产,它与组织中其他重要的资产一样需要得到适当的保护<sup>[8]</sup>。Gartner IT术语表显示:“信息(知识)资产是指与企业业务功能相关的信息,包括员工、客户或商业伙伴处获取的隐性知识;存储在高度结构化数据库里的数据和信息;以文本形式存储在非结构化数据库中的数据和信息,如电子邮件、工作流程内容和电子表格;存储在数字和纸质文档中的信息;购

买的内容以及来自互联网或其他来源的公共内容”。

从《霍利报告》给出的信息资产的定义中可以看出,信息资产是一种数据,而且是有价值的信息,包括有记录和无记录的,即任何有价值的信息,不论这个价值体现在当前还是未来。Gartner IT术语表中信息资产也被认为是一种数据或信息(有价值的信息),它可以从不同渠道获取。因此,比较合适的信息资产定义应为:已经或应该被记录的具有价值或潜在价值的信息。

## 2.2 数字资产

“数字资产”最早由Helen Meyer于1996年在《维护数字资产技巧》一文中提出<sup>[9]</sup>。2004年,Elizabeth Yakel指出“档案馆、图书馆和博物馆等已不再把数字图像等视为对象,而更多的是将其视为数字资产来对待”<sup>[10]</sup>。2006年,Albert Van Niekerk<sup>[11]</sup>给出的数字资产定义是“被格式化为二进制源代码并拥有使用权的文本或媒质等任何事物项”。2007年,Michael Moon<sup>[12]</sup>指出“可重用性”使数据文件成为资产变成可能。2011年,Laura McKinnon<sup>[13]</sup>认为在日益数字化的世界,数字资产将被作为一种新的遗产类别加以考虑和继承,在做遗产规划时需要充分了解数字资产的相关性质及变化规律。2013年,Alp Toygar等人<sup>[14]</sup>认为“从本质上说,数字资产拥有二进制形式数据所有权,是产生并存储在计算机、智能手机、数字媒体或云端等设备中的数据”。2013年,软银(SoftBank)集团的孙正义在其公司“SoftBank World 2013”活动上以“向世界挑战”为主题进行了演讲,并指出“要么数字化,要么灭亡”,认为“云将会成为人类最大的资产”。“2013英特

尔开发者论坛”上给出了包括医疗、教育、投票等在内的30个新型数字资产的类别,2015年《福布斯》杂志等媒体将比特币归为数字资产。2017年,Rod Genders和Adam Steena<sup>[15]</sup>认为“数字资产包括任何能以数字形式在线访问和持有的资产”。2018年5月,泰国颁布了《数字资产法》,包括《2018数字资产企业法》和旨在监管相关税务的《税收法修订案》两部分,法令将数字资产类型分为加密货币和数字代币(其他相似用途的电子数据单位也将被证券交易委员会授权指定为加密货币或数字令牌)<sup>①</sup>。俄罗斯等国家也将“数字资产”立法工作提上日程。

数字作为资产,被广泛关注和认可,其包含的类别在不断扩大,在一些国家已上升到立法层面,已有国家出台了相关法律,人们对其认识在不断深入。在已有数字资产的定义中,Albert Van Niekerk认为数字资产是被格式化为二进制(即被数字化)并拥有使用权的特性;Alp Toygar等人指出数字资产具有二进制形式(数字化)和所有权,并提到其产生和存储在电子设备中;Rod Genders和Adam Steena认为数字资产是数字形式和持有性。可见,二进制形式(数字化)、使用权或所有权是数字资产比较显著的特性。特别要关注的是数字资产越来越倾向于数字货币。不论是立法层面,还是当前领域研究,都将数字货币作为数字资产的重要内容来看待,更有进一步发展的趋势。从上述分析来看,比较合适的数字资产的定义应为:拥有二进制形式数据所有权,产生并存储在计算机、智能手机、数字媒体或云端等设备中的数据。

### 2.3 数据资产

“数据资产”最早由理查德·彼得

斯(Richard Peters)于1974年提出,他认为数据资产包括持有的政府债券、公司债券和实物债券等资产<sup>[16]</sup>。随着时间的推移,人们对数据资产的认识在不断深入,其内涵和范围也在不断扩展。1997年,尤谷尔·阿尔甘(Ugur Algan)<sup>[17]</sup>在《勘探生产数据库分析——实用创建技术(Anatomy of an E&P data bank: practical construction techniques)》中提到数据资产,认为“公司的市场价值和竞争定位直接关系到其数据资产的数量、质量、完整性以及由此产生的可用性,并指出创建E&P数据库是利用好数据资产的第一步”。2009年托尼·费希尔(Tony Fisher)<sup>[18]</sup>在《数据资产》中指出数据是一种资产,企业要把数据作为企业资产来对待。2009年,国际数据管理协会(DAMA)在《DAMA数据管理知识体系指南(The DAMA guide to the data management body of knowledge)》中指出在信息时代,数据被认为是一项重要的企业资产,每个企业都需要对其进行有效管理<sup>[19]</sup>。2011年,世界经济论坛(World Economic Forum)发布的《个人数据:一种新资产类别的出现(Personal data: the emergence of a new asset class)》报告中指出个人数据正成为一种新的经济“资产类别”<sup>[20]</sup>。2013年,《美国陆军信息技术应用指南(Army information technology implementation instructions)》中将数据资产定义为“任何由数据组成的实体以及由应用程序提供的读取数据的服务;数据资产可以是系统或应用程序输出的文件、数据库、文档或网页等,也可以是从数据库返回单个记录的服务和返回特定查询数据的网站;人、系统或应用程序可以创建数据资产”。2015年7月,北京中关村成立国内首家开展数据资产登记确权赋值的服务机构——

① 2018年3月13日,泰国内阁原则上批准了《2018数字资产企业法》和《税收法修订案》的草案。随后,在2018年3月27日,泰国内阁批准了两项皇家法令的最终草案,这些法令得到泰国国家委员会的审查和批准。这两项皇家法令于2018年5月13日在泰国政府公报上公布,并于次日生效

中关村数海数据资产评估中心,用于推动大数据作为资产的确权、赋值并促进交易等。2016年4月,在“全球首个数据资产评估模型发布暨中关村数据资产双创平台成立仪式”上,贵州东方世纪科技股份有限公司用数据资产进行“抵押”,拿到了贵阳银行的第一笔“数据贷”放款,中关村数海数据资产评估中心与Gartner公司一起发布了全球首个数据资产评估模型。2018年4月,中国信息通信研究院云计算与大数据研究所发布的《数据资产管理实践白皮书(2.0版)》中将数据资产定义为“由企业拥有或者控制的、能够为企业带来未来经济利益的、以物理或电子的方式记录的数据资源,如文件资料、电子数据等”<sup>[21]</sup>。

在数据资产现有的定义中,美军的定义实际上指的是与单个数据相关的系统或产品,这还不够恰当,相对而言,中国信息通信研究院给出的数据资产的定义是目前比较合适的定义。

## 2.4 概念发展

如图1所示,“数据资产”的概念最早出现,紧接着是“信息资产”的概念,直到20世纪90年代“数字资产”的概念才出现,但这3个术语的定义出现的先后顺序是:信息资产、数字资产、数据资产。

## 3 概念形成的社会背景及存在的问题

资产是经济的核心组成,一种资产类别的出现是有经济社会发展背景的。信息资产、数字资产、数据资产概念的形成与技术发展、经济社会进步相吻合。

### 3.1 信息资产概念源于信息技术发展和广泛渗透

20世纪40年代,电子计算机诞生,新技术随之出现,人们处理信息的能力得到

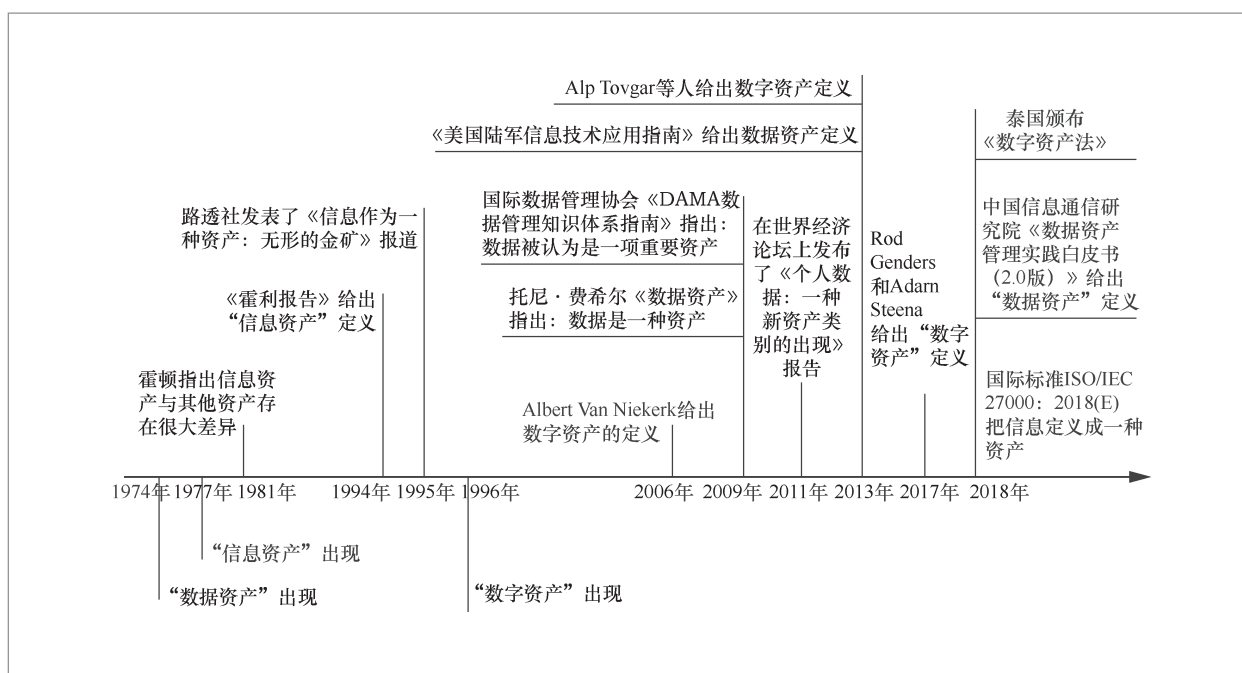


图1 数据资产相关概念发展状况

了极大的提高。1958年, Harold J. Leavitt和Thomas L. Whisler在《哈佛商业评论(Harvard Business Review)》杂志发表的文章中, 将这个新技术称为“信息技术”, 认为其由3个部分组成: 信息处理技术、统计和数学方法在决策中的应用以及通过计算机程序模拟的高阶思维<sup>[22]</sup>。信息技术迅速地改变了人类的生产和生活方式。

20世纪60年代, “信息经济”概念被提出并得以发展, 基于资源的信息观是其发展的核心, 而信息被视为资源的观点是信息资产的基础<sup>[7]</sup>。20世纪70年代中期成立的美国联邦文书委员会把信息作为一种经济商品<sup>[23]</sup>, 该委员会的信息管理研究主任霍顿认为: 信息资源管理可以有效并高效地处理信息资源及由此产生的信息资产(知识)<sup>[24]</sup>。

信息技术与各个领域的深度融合, 对各领域生产和生活方式产生了深远影响, 使经济社会的形态发生变化。在经济生产活动过程中, 信息(即有价值的信息)对于经济主体而言是非常重要的资产。

### 3.2 数字资产概念源于“数字”被广泛认识和使用

二十世纪八九十年代, 互联网得到了广泛普及, 互联网技术的快速发展推动了互联网产业化发展。相关技术的发展得益于各国实施信息高速公路计划, 从而推进了各自的信息化进程。在美国国家信息基础设施(NII)和全球信息基础设施(GII)发展的背景下<sup>②</sup>, 1998年时任美国副总统的艾伯特·戈尔在美国加州科学中心发表题为《数字地球——认识21世纪我们这个星球(The digital earth: understanding our planet in the 21st Century)》的报告, 报告中提出“数字地

球”的概念, 认为其将是“一种关于地球的可以嵌入海量地理数据的、多分辨率和三维的表示”, 并涉及包括计算机科学、海量存储、宽带网络、云数据等在内的系列关键技术, 引起了全球各界的高度关注和支持。

1995年美国麻省理工学院教授和媒体实验室前主任尼古拉斯·尼葛洛庞帝(Nicholas Negroponte)撰写的《数字化生存(Being digital)》<sup>[25]</sup>出版, 这标志着以“比特”为存在物的数字化时代已经到来<sup>[26]</sup>。

在这样的技术和时代背景下, “数字经济(Digital economy)”概念被提出并迅速流行, 加快发展数字经济已成为各国共识。在数字经济发展过程中, 数字资产被日益关注。

### 3.3 数据资产概念源于大数据的兴起

1997年NASA研究员Michael Cox和David Ellsworth在IEEE第8届国际可视化学术会议上首先提出了“大数据”术语<sup>[27]</sup>。2008年9月《自然》杂志出版了一期大数据专刊, 使得大数据在科学研究领域得到了高度重视。2012年3月美国政府发布了《大数据研究和发展倡议(Big data research and development initiative)》, 大数据引起了很多国家和全社会的重视<sup>[28]</sup>。

大数据时代, 人们的关注点是数据。数据作为一种重要的战略资源, 其价值被不断挖掘和创造, 数据正在成为一种新的资产——数据资产。在大数据时代, 数据是人们关注的资源, 能让机器读取的数据才有可能读出有价值的信息, 成为数据资产。

### 3.4 存在的问题

经济社会历史发展环境的不同和对

②

1993年9月, 美国政府发表《国家信息基础设施: 行动计划》, 建设“信息高速公路”, 实现大量信息共享。1994年3月, 美国前副总统艾伯特·戈尔在国际电信联盟(ITU)主持召开的首届世界电信发展会议上, 提出并号召各国积极参与建设全球信息基础设施的宏伟构想, 并得到各国广泛响应

“信息、数字、数据”概念认识的不同,形成了不同的名词术语,但3个概念名词的内涵差异不大,在具体使用时也会出现混乱和相互替代的情形。更为严重的是,资产和资源、资本、经济等术语紧密关联,于是就有了信息资产、信息资源、信息资本、信息经济、数据资产、数据资源、数据资本、数字经济、数字资产、数字资源、数字资本和数字经济12个概念。

显然,过多的、雷同的概念不利于事情的健康发展。信息资产、数字资产、数据资产3个概念本质上都是在讲数据。在信息资产的一些使用场合,会把存放在纸质材料的数据也作为信息资产,例如:国际标准ISO/IEC 27000:2018(E)中指出信息能以多种形式存储,包括数字形式(如电子或光媒质上存储的数据文件)、物质形式(如纸质形式)以及员工的知识形式,信息可以通过各种方式传递,包括信使、电子或口头交流<sup>[8]</sup>。显然,纸质形式和电子形式在规模、流通方式上存在本质区别,例如:1 PB的数据相当于30个国家图书馆的藏书规模,而1 PB规模是大数据的一个基础度量。因此,需要将数据资源界定在网络空间中,不考虑网络空间外的数据资源。

网络空间是指计算机网络、广电网、通信网络、物联网、卫星网等所有人造网络和设备构成的空间,这个空间真实存在。电脑、手机、移动硬盘等都是网络空间的组成部分。现在,空调、冰箱、自动窗帘、电子门锁等也已经成为网络空间的组成部分。数据是指网络空间的任何事物,“数据”的含义很广,不仅包括传统意义上的数据(如1011、8084等),还包括符号、字符、日期形式的数据(如dataology、上海市数据科学重点实验室、2013/09/06等),也包括文本、声音、图像、照片和视频等类型的数据,微博、微信、购物记录、住宿记

录、乘坐飞机记录、银行消费记录、政府文件等也是数据。

2009年,朱扬勇、熊贇提出数据有物理属性、存在属性和信息属性<sup>[29-30]</sup>。

- 物理属性:是指数据在存储介质中以二进制形式存在。数据的物理存在确实占用了存储介质的物理空间,是数据真实存在的表现,并且可以度量。数据的物理存在可以直接用于制作数据复本和数据传输。

- 存在属性:是指数据以人类可感知(通常为可见、可听)的形式存在。在计算机系统中,物理存在的数据需要通过I/O设备以日常的形式展现出来,才可以被人感知、认识。人通过I/O设备感知到的(比如看见的)数据被认为是在数据界中存在的

- 信息属性:一个数据是否有含义?含义是什么?这是数据的信息属性。通常数据经过解释就会有含义,数据的含义就是信息;也有一些数据是没有含义的,例如,一个随意打入的字符串“20 x ad7f海7s9f9a-877w2”就没有含义,但它是数据自然界中的一个数据。

数据的物理属性是数据的数字形态、物理形态,是有形的;数据的存在属性是数据的感知性,数据在存在属性上能够被人感知到,如果感知不到,就可以认为数据不存在。因此,数据的物理属性和存在属性是相互依附的,物理上存在的数据必须能够被感知到,感知到存在的数据必须在物理上存在。

数据的信息属性其实就是数据的价值体现,数据可以没有价值,但作为资产的数据必须有价值。

从以上分析可以看到,信息资产、数字资产、数据资产是从不同层面看待数据的。信息资产对应着数据的信息属性,数字资产对应着数据的物理属性,数据资产

对应着数据的存在属性。图2展现了4个数据集在物理属性、存在属性和信息属性上的内容。从目前的定义来看,它们都是数字资产和数据资产,但数据集4不是信息资产。

## 4 定义数据资产

根据第3.4节的分析,有必要将信息资产、数字资产和数据资产统一为数据资产,并依照数据属性开展数据的资产化研究工作。

### 4.1 数据具备资产的基本条件

资产是指会计主体(政府、企事业单位等)由过去的经济业务或者事项形成的、由会计主体控制的、预期能够带来经济利益流入或产生服务潜力的经济资源。资产的分类很多,如流动资产、固定资产、有形资产、无形资产、不动产等。

资产具有以下几个方面的特征:

- 资产预期会给会计主体带来经济利益或产生服务潜力;
- 资产应是会计主体拥有或者控制的资源;
- 资产是由会计主体过去的交易或者事项形成的。

按照中国的企业会计准则,符合上述资产定义的资源还要同时满足以下条件,才能被确认为资产:

- 与该资源有关的经济利益很可能流入企业;
- 该资源的成本或者价值能够可靠地计量。

在大数据背景下,数据资源是重要的战略资源已成为共识<sup>[31]</sup>。数据是当前社会、经济运行过程中不可或缺的部分。那么,数据是否具备了资产的特征和条件呢?

首先,数据能够给会计主体带来利益。在经济业务活动中,会计主体可通过生产、采集、加工、购买等方式拥有并控制数据资源,并且可以通过出让数据、加工数据或提供数据服务来获取利益。数据资源是能给会计主体带来预期经济利益或产生

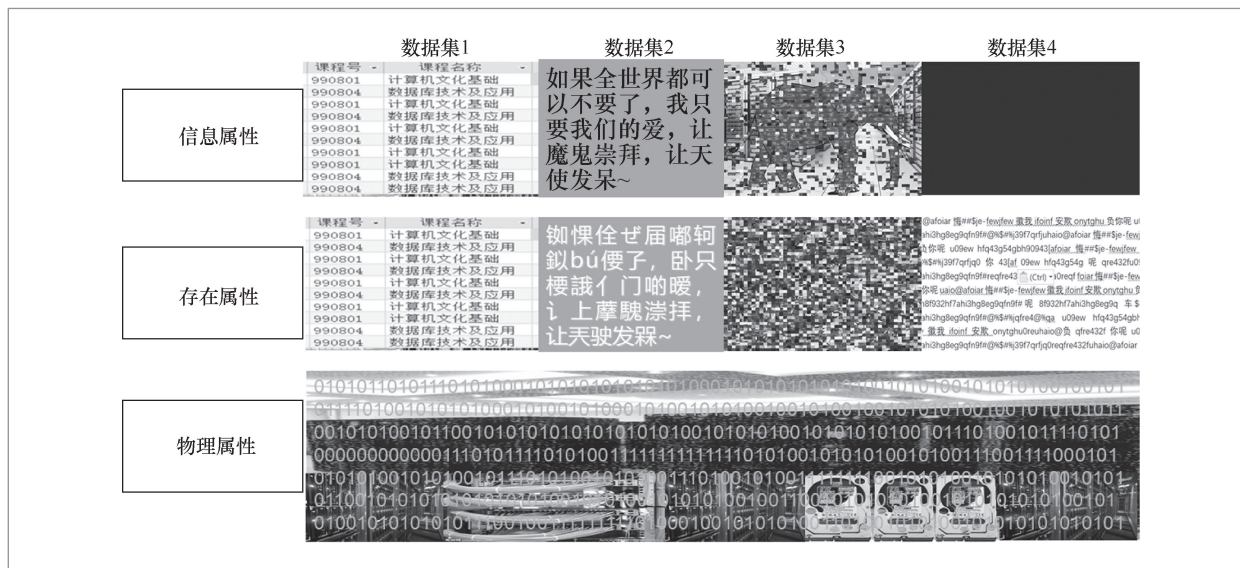


图2 4个数据集的属性

服务潜力的资源。当然,并不是所有的数据都能够给会计主体带来利益,垃圾数据是不会带来利益的。

其次,数据的成本或者价值能够被可靠地计量。现有的定义问题在于“计量”,由于计量问题,数据资产迟迟难以进入会计报表。由于数据是现实事务的信息化,因此非常复杂,难以用一个统一的计量方法对各类数据进行计量。目前,已经有一些数据可以可靠地计量了,例如,2003年乔布斯曾经创造性地重构了音乐的价格机制,苹果iTunes音乐商店中的所有音乐,无论作者、长短、受欢迎程度,统一定价为99美分,开创了以单首音乐为计量单位、99美分为计价的新型定价模式。这样,音乐作为一种数据资产就得以实现。

由此可见,能够带来利益的数据具备了资产的基本特性,如果能够计量,就符合了资产的条件。不能够带来利益的数据显然不是资产,不能够被可靠计量的数据不能被计入会计报表,暂不适合作为资产来看待。

## 4.2 数据资产的定义

大数据通常指一个大规模(常常是PB级别)的数据集,需要通过开发这个数据集获得价值,而1 PB规模的数据集就相当于30个国家图书馆的数据量。因此,区别于传统的“数据”概念,本文将“数据”界定为网络空间的任何事物(不包括纸质或其他非网络空间的数据),它在网络空间中是唯一存在的。结合数据属性,将数据资产定义为:拥有数据权属(勘探权、使用权、所有权)、有价值、可计量、可读取的网络空间中的数据集。

首先,因为纸质资料和电子资料在计量和规模上有着本质的不同,所以本定义将数据界定为网络空间中的数据,以排除

图书馆、档案馆的纸质资料数据资产。除了涉及非网络空间的数据之外,在将数据界定为网络空间中数据的前提下,该定义与现有的关于信息资产、数字资产和数据资产的主要定义之间不存在不一致的问题。

其次,主要的差别是“可计量”。无论是无形资产还是有形资产,可计量是资产化的必要条件,只有可计量才有可能进入会计报表,资产化才可能实现。数据是现实世界中万事万物的信息化,数据类型、规格、内容也多种多样,有些可以计量、有些目前还没有找到计量方法。先把能够计量的数据资产化是一个可行的做法,因此,把可计量纳入数据资产的定义中。

再次,关于数据权属应考虑勘探权、使用权、所有权。不同于矿山、矿藏,它们的所有权属于国家,而数据是非天然的,情理上数据应属于数据的生产者。企业、机构和个人可以持有数据的所有权。

最后,根据定义可知,下列数据集不是数据资产:

- 没有价值的数据集、垃圾数据集;
- 没有数据权属的数据集;
- 不能计量的数据集;
- 不可读的数据集。

## 4.3 数据资产是一种新的资产类别

和数据一样,数据资产也有物理属性、存在属性和信息属性。

● 数据资产的物理属性是指数据资产在存储介质中以二进制形式存在,占有物理空间,是有形的。在大数据背景下,1 PB的数据集是经常出现的,通常1 PB的数据集会占用3 PB的存储空间,就像存放石油需要一个油库一样,需要为其准备一个有实物形态的仓库(如物理存储设备、机房等)来保存,在物理上是有形的。

- 数据资产的存在属性是指可读取性。不可读取意味着资产不可见，价值就不会实现。若将石油装在桶里，是可见的，是有形的，若将一桶石油洒在地上，看不见了、没了，就不是资产了。因此，数据资产的物理属性加上存在属性就形成了数据资产的物理存在，是有形资产。

- 数据资产的信息属性是其价值所在。由于信息及其价值都难以计量，因此信息资产自提出以来，就被纳入无形资产范畴。事实上，一个数据资产包含的信息的价值取决于挖掘、使用信息的人，几乎不可能有统一的价值标准。因此，对数据资产包含的信息价值只能采用评估的方式，从这个角度来讲，数据资产属于无形资产。

由此可见，数据资产的物理属性和存在属性表现出有形资产的特征，而其信息属性以及数据勘探权、使用权等则表现出无形资产的特征，即数据资产兼有无形资产和有形资产的特征。另外，数据的复制成本远低于生产成本，因此数据的流动性极好，可以在一个会计年度随意流通，具备流动资产的特征；同时数据的使用没有损耗，不具有实物资产的折旧，数据可以长期存在，因此又具备长期资产的特征。在会计领域，需要将数据资产作为一种新的资产类别予以单列。

## 5 结束语

本文综述了信息资产、数字资产、数据资产的概念形成和发展过程，在大数据背景下，将“数据”界定为网络空间中的事物，不包括纸质资料等非网络空间的数据形式，依据数据的物理属性、存在属性和信息属性，将信息资产、数字资产、数据资产3个概念统一成数据资产。重新将数据

资产定义为拥有数据权属（勘探权、使用权、所有权）、有价值、可计量、可读取的网络空间中的数据集。给出了数据资产的性质，为后续数据资产的会计处理打下了基础。后续工作主要包括以下几个方面。

- 研究数据资产权属：从法律方面来看，相对于数据生产的成本来说，数据复制的成本极低，数据的复制传播非常容易，因此要研究数据使用权对数据的所有权造成伤害的问题，合理界定数据的勘探权、使用权、所有权。从会计方面来讲，一个数据集的勘探权和使用权应归为无形资产；一个数据集的所有权应归为有形资产。

- 建立数据资产的计量计价方法：从目前音乐、电影、小说等数据产品的销售实践来看，以数据资产的物理属性来建立数据资产的计量计价方法是可行的。需要研究计量计价的可行性、基本要求和相应标准。从会计方面来讲，能够计量的数据资产可先行纳入会计报表。

- 研究数据资产会计类别和处理方法：数据资产兼有无形资产和有形资产、流动资产和长期资产的特征，是一种新的资产类别，需要专门研究数据资产的会计处理方式。

## 参考文献：

- [1] KABACK S M. A user's experience with the derwent patent files[J]. Journal of Chemical Information and Computer Sciences, 1977, 17(3): 143-148.
- [2] HORTTON F W. Information resources management (IRM): where did it come from and where is it going[C]// The ACM' 81 Conference, April 29-May 1, 1981, Ann Arbor, UAS. New York: ACM Press, 1981: 277-278.
- [3] KPMG/IMPACT. Information as an asset: the board agenda[R]. London: KPMG/

- IMPACT Group, 1994.
- [4] WADDINGTON P. Information as an asset: the invisible goldmine[J]. *Business Information Review*, 1995, 12(1): 26-36.
- [5] HAWLEY R. Information as an asset: the board agenda[J]. *Information Management and Technology*, 1995, 28(6): 237-239.
- [6] OPPENHEIM C. Valuing information assets in British companies[J]. *Business Information Review*, 1998, 15(4): 209-214.
- [7] STENSON J. Senior UK managers' identification of attributes of information assets[J]. *Advances in Library Administration and Organization*, 2006(23): 333-414.
- [8] ISO/IEC27000:2018(E). *Information technology-Security techniques-Information security management systems-overview and vocabulary*[S]. 2018.
- [9] MEYER H. Tips for safeguarding your digital assets[J]. *Computers & Security*, 1996, 15(7): 588.
- [10] YAKEL E. Digital assets for the next millennium[J]. *Oclc Systems & Services*, 2013, 20(20): 102-105.
- [11] NIEKERK A V. A methodological approach to modern digital asset management: an empirical study [C]// *Allied Academies International Conference, International Academy for Case Studies*, January 1, 2006, New Orleans, USA. [S.l.:s.n.], 2006.
- [12] MOON M. Cycle time activity lifecycle of digital assets[J]. *Journal of Digital Asset Management*, 2007, 3(3): 112-115.
- [13] MCKINNON L. Planning for the succession of digital assets[J]. *Computer Law & Security Review*, 2011, 27(4): 362-367.
- [14] TOYGAR A, ROHM C E T J, ZHU J. A new asset type: digital assets[J]. *Journal of International Technology & Information Management*, 2013, 22(4): 22.
- [15] GENDERS R, STEENA A. Financial and estate planning in the age of digital assets: a challenge for advisors and administrators[J]. *Financial Planning Research Journal*, 2017, 3(1): 75-80.
- [16] PETERSON R E. A cross section study of the demand for money: the United States, 1960-62[J]. *Journal of Finance*, 1974, 29(1): 73-88.
- [17] ALGAN U. Anatomy of an E&P data bank: practical construction techniques[J]. *The Leading Edge*, 1997, 16(6): 901-902.
- [18] FISHER T. *The data asset: how smart companies govern their data for business success*[M]. New York: Wiley, 2009.
- [19] DMMA International. *The DAMA guide to the data management body of knowledge-DAMA-DMBOK*[M]. [S.l.]: Technics Publications, 2009.
- [20] MICHELE. *Personal data: the emergence of a new asset class*[R]. 2011.
- [21] 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(2.0版)*[R]. 北京: 中国信息通信研究院, 2018.
- Cloud Computing and Big Data Research Institute of China Academy of Information and Communications Technology. *White paper of data asset management from the perspective of practice*[R]. Beijing: China Academy of Information and Communications Technology, 2018.
- [22] LEAVITT H J, WHISLER T L. *Management in the 1980's*[J]. *Harvard Business Review*, 1958.
- [23] HORTON F W. *Information resources management: concept and cases*[M]. Cleveland: Association for Systems Management, 1979.
- [24] HORTON F W. *Information resources management: harnessing information assets for productivity gains in the office, factory and laboratory*[M]. Englewood Cliffs: Prentice Hall, 1985.
- [25] NEGROPONTE N. *Being digital*[M]. New York: Alfred A. Knopf, 1995.
- [26] 蔡曙山. 论数字化[J]. *中国社会科学*, 2001(4): 33-42.
- CAI S S. On digitalization[J]. *Social Sciences in China*, 2001(4): 33-42.
- [27] COX M, ELLSWORTH D. *Application-controlled demand paging for out-of-*

- core visualization[C]//The 8th Conference on Visualization, October 18-24, 1997, Phoenix, USA. Los Alamitos: IEEE Society Press, 1997: 235-244.
- [28] 朱扬勇, 熊贇. 大数据是数据、技术, 还是应用[J]. 大数据, 2015,1(1): 71-81.  
ZHU Y Y, XIONG Y. Defining big data[J]. Big Data Research, 2015, 1(1): 71-81.
- [29] ZHU Y Y, ZHONG N, XIONG Y. Data explosion, data nature and dataology [C]//The 2009 International Conference on Brain Informatics, October 22-24, 2009, Beijing, China. Heidelberg: Springer, 2009: 147-158.
- [30] 朱扬勇, 熊贇. 数据学[M]. 上海: 复旦大学出版社, 2009.  
ZHU Y Y, XIONG Y. Dataology and data science[M]. Shanghai: Fudan University Press, 2009.
- [31] 朱扬勇. 大数据资源[M]. 上海: 上海科学技术出版社, 2018.  
ZHU Y Y. Data resource[M]. Shanghai: Shanghai Scientific & Technical Publishers, 2018.

### 作者简介



**朱扬勇** (1963-), 男, 博士, 复旦大学计算机科学技术学院教授、学术委员会主任, 上海市数据科学重点实验室主任。1989年起从事数据领域研究, 2008年提出数据资源保护和利用, 2009年发表了数据科学论文“Data explosion, data nature and dataology”, 并出版专著《数据学》, 对数据科学进行了系统探讨和描述。2010年创办了“International Workshop on Dataology and Data Science”, 2014年和石勇、张成奇共同创办了“International Conference on Data Science”。担任第462次香山科学会议“数据科学与大数据的理论问题探索”的执行主席、《大数据技术与应用丛书》主编。目前主要研究方向为数据科学、大数据。



**叶雅珍** (1985-), 女, 复旦大学计算机科学技术学院助理研究员, 东华大学博士生。2014年起从事数据领域方面的研究工作, 作为项目主要人员参与过多项国家自然科学基金项目和与上海市科学技术委员会、地方政府单位以及企业合作的项目, 并发表论文多篇。主要研究方向为数据科学、大数据、数据资产。

收稿日期: 2018-09-29

基金项目: 上海市科技发展基金资助项目 (No. 16JC1400801)

Foundation Item: Shanghai Science and Technology Development Fund (No. 16JC1400801)