

研究报告

2020 年第 60 期

2020.12.18

执笔人：朱俊丰

junfeng.zhu@icbc.com.cn

数据要素的发展与启示

要点：

- 数据的本质是信息的表现形式和载体，数据要素具有边际成本递减、非竞争性、非排他性、外部性等经济特征以及偏向性技术进步、非结构性、非标准化、集合化、多样性等技术特征。
- 数据发展呈现出数据量与市场规模快速增长的态势。未来，数据要素将会向以公有云存储数据为主导，实时数据需求驱动，数据阵地由端点向核心、边缘转移的趋势发展。
- 我行发展数据要素的对策与建议：一是立足数据要素管理的发展战略，二是实施数据存储“自建”与“合作”并行的发展模式，三是加强数据要素的分析与应用能力，四是完善数据资源的安全保护水平。

重要声明：本报告中的原始数据来源于官方统计机构和市场研究机构已公开的资料，但不保证所载信息的准确性和完整性。本报告不代表研究人员所在机构的观点和意见，不构成对阅读者的任何投资建议。本报告（含标识和宣传语）的版权为中国工商银行现代金融研究院所有，仅供内部参阅，未经作者书面许可，任何机构和个人不得以任何形式翻版、复制、刊登、上网、引用或向其他人分发。

随着新一轮信息技术的迅猛发展，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》明确将数字与土地、劳动力、资本、技术并列为五大生产要素，以大数据、物联网、云计算、人工智能为基础的数字经济正逐渐成为国家经济发展的重要动能。随着生产要素范围的拓展，生产效率与资源配置将会面临新的变化与调整，经济发展也将由基于传统生产要素的发展模式转向基于数字化知识与信息利用的高质量发展模式，企业的发展战略与管理模式也随之面临新的挑战与机遇。在当今世界百年未有之大变局下，中国亟需立足数据要素的赋能机遇，识别数据要素的基本特征及问题，加快培育数据要素市场，推动企业数字化转型，促进实体经济和数字经济融合嬗变，保障中国经济的稳定与高质量发展。

一、数据要素的定义与基本特征

（一）数据要素的定义

数据本质上是对物品、服务或经济主体等相关信息的电子或非电子形式的记录，是信息的表现形式和载体。一般而言，日常的数据主要包括四类：**一是**最终商品或服务数据，如在线咨询、视频等；**二是**直接进行交易的数据，如大数据交易所的交易数据；**三是**帮助提升最终产品或服务的性能或生产效率的数据，如引流、效果广告、配送优化等；**四是**在兼并收购或战略合作中有价值体现的数据，是并购或合



作谈判筹码，并非直接交易数据¹。

数据并非天然就是生产要素。一般而言，日常接触的数据主要是最终产品或服务的数据，由于其不参与生产加工的流程，因此并非生产要素，而第二类至第四类的数据才是生产要素数据。

表 1 数据的类型与内涵界定

	类型	示例
普通数据	最终商品或服务数据	在线咨询、视频
生产要素	直接进行交易的数据	交易所数据
	帮助提升最终产品或服务的性能或生产效率的数据	引流、效果广告、配送优化
	在兼并收购或战略合作中价值体现的数据	并购或合作谈判数据

资料来源：吴绪亮. 制度设计是数据要素定价的关键[N]. 经济日报，2020-10-23(011)。

（二）数据要素的基本特征

一方面，数据要素具有边际成本递减、非竞争性、非排他性、外部性等经济特征²。一是边际成本递减效应，数据一旦被产生且被收集打包，任何人均可使用，且边际成本几乎为零。二是数据具有非竞争性与非排他性特征，共享程度较高，但其所有权归属较难明晰，难以有效参与市场化交易。三是数据要素作为一种公共产品或准公共产品，本身极易产生外部性，如数据精准匹配产生经济效益的正外部性

¹ 吴绪亮. 制度设计是数据要素定价的关键[N]. 经济日报，2020-10-23(011).

² 中国人民大学与中信证券研究部课题组. 金融科技创新发展研究报告：数据要素与金融科技创新[R]，2020。

以及隐私泄露所产生的数据安全负外部性等问题。

另一方面，数据要素具有偏向性技术进步、非结构性、非标准化、集合化、多样性等主要技术特征¹。一是以人工智能、大数据、云计算为代表的数字技术主要应用于信息通信技术产业、金融、电商平台、互联网等领域，表现出显著的偏向性技术进步特征。二是非结构化与半结构化程度明显，80%的全球信息数据是以文件形式存在的非结构化和半结构化数据。三是非标准化问题突出，术语标准、元数据和标识标准、质量标准以及资产管理标准等问题亟待解决。四是集合数据的“大数量、多类型、价值密度低、高时效”等综合特征使得数据成为新型生产要素。五是大数据多样性既涵盖了文本、音频、视频、图像、模拟信号等格式多样性，又包括了区域、行业、企业、个人等来源多样性。

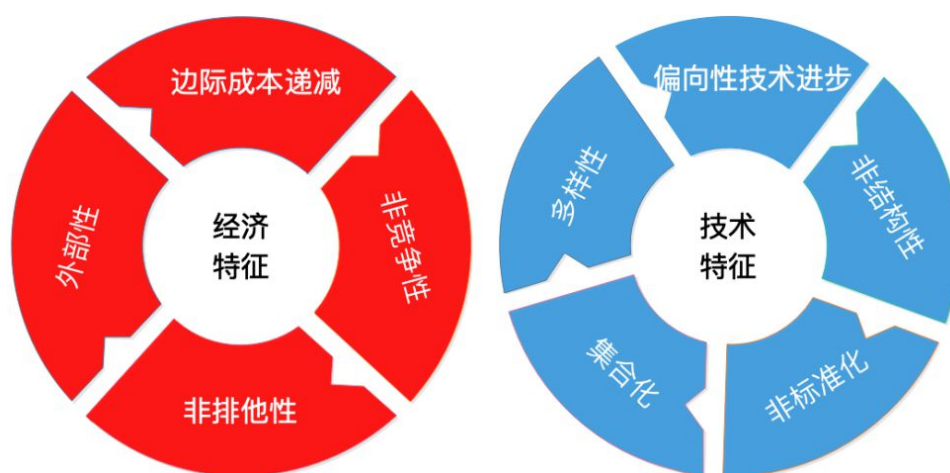


图1 数据要素的经济特征与技术特征

¹ 绉传伟，对数据要素的特征、价值和配置机制的初步研究[EB/OL].<https://www.sohou.com/a/393564350-100217347>, 2020-05-07。



二、数据要素的发展现状与趋势

（一）数据要素的发展现状

1.全球数据量与中国数据量呈现爆发式增长。IDC 的研究报告显示，新摩尔定律主导下全球数据量将从 2019 年的 45ZB 增长到 2025 年的 175ZB，呈现出增长率逐年提高的级数增长态势。中国的数据总量将从 2018 年的 7.6ZB 增长至 2025 年的 48.6ZB，占全球数据总量的比重将会由 23%提升至 27.8%。

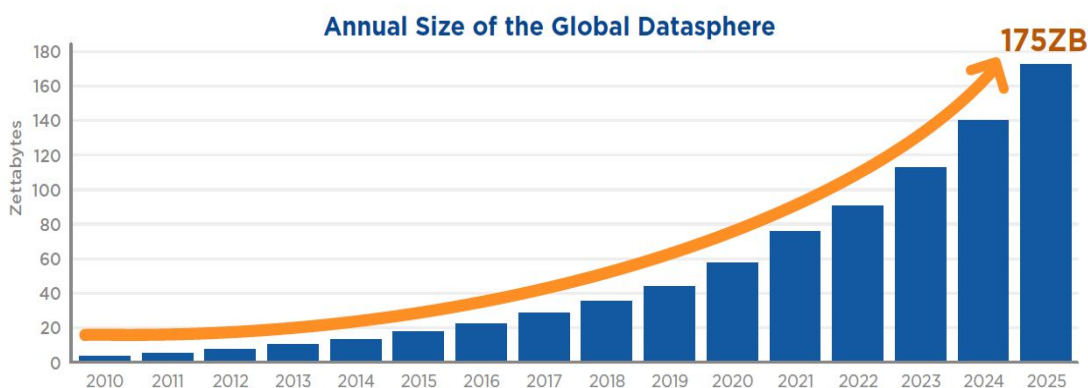


图2 全球数据量的发展及预测情况

资料来源：Data Age 2025, sponsored by Seagate with data from IDC Global Datasphere, May 2020.

2.中国数据要素市场规模快速增长。根据中国信通院发布的《2020 年中国数字经济发展白皮书》数据显示，十三五期间，以数据采集、数据清洗、数据标注、数据交易等核心数据要素环节构成的中国数据要素市场规模快速增长。预计在 2020 年将达 545 亿元，十三五期间数据要素市场复合增速超过 30%。

（二）数据要素的发展趋势

1.公有云存储数据将占据主导地位。随着更高智能水平的设备连接,端点设备存储容量较小的问题使得数据管理逐渐向企业设备和公有云转变。根据 IDG 的研究分析,在 2020 年,公共云上存储的数据量所占比重将会超过消费设备上存储的数据占比,到 2022 年,将会有更多的数据储存在公共云上而非企业设备,具体如图 3 所示。

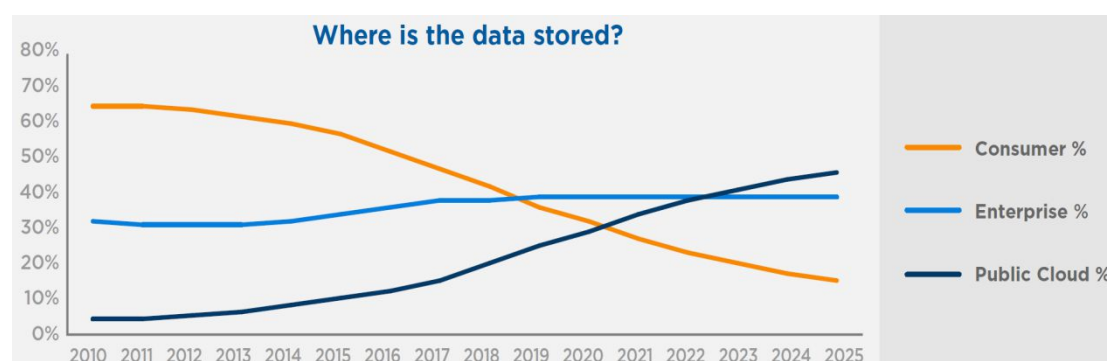


图 3 数据存储中心的变化情况

2.实时数据需求驱动。目前,每天有超过 50 亿消费者与数据互动,平均每人每天会进行 1426 次的数字数据参与。到 2025 年,互动人数将达到 60 亿,平均每人每天会进行 4909 次的数字数据参与。根据 IDC 预测,2025 年将会产生超过 90ZB 的实时交互数据,占全球数据圈的近 30%。

3.数据阵地将由端点向核心、边缘转移。从全球数据创建的角度来看,整体呈现出端点创建数据比例逐渐下降,核心和边缘的创建数据比例逐年递增的发展趋势。对中国而言,目前近 60%的数据在端点创建,但越来越多的数据将会在边缘被复制、传输或备份,且最终



将在核心复制¹。

三、对我行的启示

一是立足数据要素管理的战略层面。提升数字思维能力，积极响应国家数据要素改革任务，结合我行数字化转型目标，将数据驱动型创新作为数字化转型战略的重中之重。制定数据管理与信息治理策略，总体把控我行未来数据要素发展的方向、重点任务及数据治理体系。

二是实施数据存储“自建”与“合作”并行的模式。一方面，利用云平台存储数据，加快单一的公有云模式向多云、混合云基础架构转变。另一方面，在已有企业级数据中台的基础上进一步优化平台的开放、共享、共建功能，强化数据获取、集成整合、挖掘分析、即时赋能等核心功能。

三是加强数据要素的分析与应用能力。优化企业级技术架构与大数据分析能力，加快数据要素向数据资产转变；强化我行在端点、核心和边缘之间循环传播数据的相互作用，以应用于欺诈检测等情况；深挖数据要素的融合价值，强化数据协同与数据复用，重视高频数据，积极探索政企、企企数据的开放、共享与合作；重视对数据分析师等专业人才的培养与挖掘，丰富数据应用手段。

¹ 端点是指网络边缘中的所有设备，包括电脑、电话、工业传感器、联网汽车和可穿戴设备；核心是指企业和云提供商专门的计算数据中心；边缘是指由企业强化、不位于核心数据中心的服务器和设备。

四是完善数据资源的安全保护水平。一方面优化数据隐私保护，完善适用于大数据环境下的覆盖数据全生命周期的分类分级安全保护标准；另一方面，强化数据安全审查，明确数据资产管理责任主体。