

数据要素与先进存储 融合发展研究报告

中国信息通信研究院产业与规划研究所

2024年9月

版权声明

本报告版权属于中国信息通信研究院，并受法律保护。转载、摘编或利用其它方式使用本报告文字或者观点的，应注明“来源：中国信息通信研究院”。违反上述声明者，本院将追究其相关法律责任。

前 言

数据作为新型生产要素，是基础性和战略性资源，也是发展新质生产力的重要基础。世界各国高度重视数据价值释放，争相抢占数据资源新高地，塑造国家竞争新优势。我国率先将数据列入生产要素，率先构建数据交易场所，率先开展数据资产入表，引领全球数据要素市场发展潮流。

新技术、新模式、新基建带来数据要素领域的新发展。以人工智能为代表的新技术加快数据要素价值释放。随着人工智能大模型的到来，冷数据逐渐变热，也产生了更多热数据。大模型训练数据源深度持续拓展，数据存储周期显著加长。人工智能大模型对高质量数据集也提出了新要求，数据集的质量影响人工智能的智商，人工智能大模型的训练和推理越来越依赖高质量数据集的提供和先进存储的支撑。

数据资产化带来“数据信贷”“数据信托”等数据要素价值化新模式的探索，为数据要素价值释放创造了基础条件。数据要素促进金融市场发展，数据抵押贷款、数据信托等金融产品反过来又拓展了数据要素的应用场景。数据资产流通交易也带来了数据副本量的快速增加。数据已成为数字经济时代最重要的资产之一。正如金融资产存在银行保险箱里一样，数据资产也需要有安全可靠的保险库。先进存储设施正是存储数据资产的保险库，是存储数据资产的最重要的基础设施，为数据的资产化、产权化提供坚强保障。

数据基础设施是典型的新基建，与数据要素如影随形、相生相伴、相辅相成，但二者也是矛盾的关系。数据基础设施是数据“供得出、流得动、用得好、保安全”的关键载体。数据基础设施的建设为数据的全生命周期管理构筑了坚实支撑。存储设施与网络、算力、数据采集、数据流通、数据安全等设施共同构成数据基础设施的“六个底座”，

是“六位一体”的关系。特别是，数据和存储设施的关系就像水和瓶子一样，随着水流量的大规模增加，需要更大容量的瓶子，且这个瓶子不能存在瓶颈。“瓶能装水亦能覆水”，“覆水难收”，瓶子要坚固。同时，也要尽可能实现节能环保。反之，随着瓶子容量、质量的提高，会带动更大规模水流量的产生，对瓶子的要求又进一步提高。如此往复，呈螺旋上升之势。

为进一步促进数据要素领域发展，应建设高质量数据集，强化新型人工智能存储创新与应用；分步推动数据资产入表，优化数据资产评估体系；加快建设先进存储设施，优化我国数据基础设施建设布局，将数据基础设施打造成促进数据要素发展的坚强底座。

目 录

一、数据要素发展势头强劲，成为经济社会新动能.....	1
（一）我国数据要素领域面临重大发展机遇	1
（二）我国数据存储领域呈现快速发展态势	3
（三）新技术新模式新基建促进数据要素新发展	5
二、数据要素与人工智能相互促进，先进存储奠定坚实基础.....	7
（一）人工智能加速数据要素升温，冷数据向温热数据转变	7
（二）人工智能提升数据存储价值，数据保存周期显著加长	8
（三）高质量数据集助力人工智能，先进存储加速数据归集	9
（四）人工智能对存储提出高要求，新型人工智能存储取得突破	10
三、数据资产化推动数据要素价值释放，先进存储成为数据资产保险库.....	11
（一）数据资产评估入表和产权登记取得重要突破	12
（二）数据资产流通交易带来数据副本量快速增加	15
（三）先进存储为数据资产提供安全可靠的保险库	16
四、先进存储设施是数据基础设施的关键一环.....	17
（一）数据基础设施是关键性载体，存储是“六个底座”之一	17
（二）数据基础设施建设提速提质，先进存储占比不断提升	19
（三）存力中心建设成为典型模式，各地建设取得重要成效	22
五、发展建议.....	24
（一）推进高质量数据集建设，强化新型人工智能存储创新应用	24
（二）完善数据资产评估体系，增加数据资产的存储安全性评估	24
（三）加快建设先进存储设施，优化我国数据基础设施建设布局	25

图 目 录

图 1 数据使用范围与对应的数据副本量增加情况.....	16
图 2 数据基础设施组成部分	19
图 3 2019 年-2023 年我国数据中心机架数.....	21
图 4 2021 年-2025 年我国存储总量及先进存储占比.....	22

CAICT 中国信通院

党的十九届四中全会将数据列入生产要素，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”。这是全球范围内首次将“数据”明确作为生产要素，是巨大的理论创新，具有里程碑意义。数据作为新型生产要素，与各类新技术、新模式、新基建紧密相关、相辅相成。特别是，数据要素与数据基础设施相互依存、相互促进。其中，作为数据基础设施重要组成部分，先进存储设施代表存储领域未来发展方向，具有大容量、高效率、高可靠和绿色低碳等特点。先进存储具有的这些优势为数据要素发展奠定了坚实基础。所谓**先进存储**是指应用全闪存阵列、SSD等先进存储部件，采用存算分离、高密等先进技术，单位容量数据操作能力达到万 IOPS（每秒读写次数）以上的存储模块¹。

一、数据要素发展势头强劲，成为经济社会新动能

（一）我国数据要素领域面临重大发展机遇

各国积极推动数据价值释放，数据成为国家基础性战略性资源。当前，全球数据量井喷式增长，数据已融入生产生活各方面，深刻影响并重构着经济社会运行和社会治理，成为影响国家未来发展的关键资源。据 IDC 预测，2024 年全球将生成 159.2ZB 数据，2028 年将增加一倍以上，达到 384.6ZB，复合增长率为 24.4%。全球各国很早就认识到数据的战略地位。美国最先发布《大数据研究和发展倡议》，引发全球大数据发展浪潮，并以“政府引导、企业参与、市场运作”的模式为主，投入大量资金用于数据基础设施建设。欧盟率先发布数据战略，并于 2022 年 4 月通过《数据治理法案》，作为落实战略的具体举措。为充分发挥海量数据对经济发展的促进作用，提高数据应用

¹来源：《工业和信息化部等六部门关于印发<算力基础设施高质量发展行动计划>的通知》

水平，释放数据要素价值，各国围绕推进数据基础设施建设、构建数据可信流通环境、完善数据法律法规、加强数据领域国际合作等方面，全方位加快布局建设，争相抢占数据资源新高地，塑造国家竞争新优势。

我国数据要素政策持续出台，数据基础制度体系不断完善。我国高度重视数据领域高质量发展，出台了一系列推动数据要素体系化建设的政策文件，为推动数据要素价值释放提供了指引。2022年12月，党中央、国务院印发《关于构建数据基础制度更好发挥数据要素作用的意见》（以下简称“数据二十条”），系统性布局数据基础制度体系的“四梁八柱”，绘制数据要素发展的长远蓝图。为贯彻落实“数据二十条”，国家数据局等十七个部门联合印发《“数据要素×”三年行动计划（2024—2026年）》，提出数据要素发展总体目标和十二项重点行动，推动数据要素发挥乘数效应，赋能经济社会发展。财政部出台《关于加强数据资产管理的指导意见》，明确了依法依规管理数据资产、明晰数据资产权责关系、完善数据资产相关标准、加强数据资产使用管理、稳妥推动数据资产开发利用、健全数据资产价值评估体系等12个方面的主要任务。我国数据基础制度体系加快推进建设，统筹管理、协调发展的体制机制进一步完善。

各地区数据管理机构纷纷设立，数据工作体系初步形成。为深入贯彻党中央和国务院关于数据工作的决策部署，2023年10月，国家数据局正式挂牌成立，负责协调推进数据基础制度建设，统筹数据资源整合共享和开发利用，统筹推进数字中国、数字经济、数字社会规划和建设等。从地方层面看，近些年来，31个省（区、市）和新疆生产建设兵团陆续完成数据管理机构的组建工作，初步形成了上下联动、横向协同的全国数据工作体系。数据管理体制机制的建设加快推动了

数据基础制度构建，是抢抓数字经济发展先机、打造经济发展新动能的重要举措。

各行业持续推动数字化转型，数据要素乘数效应加速显现。数据要素领域建设的核心目标是深化数据融合应用，推动数据要素进一步赋能，强化数据在提高生产效率与资源配置效率方面的作用，培育新质生产力。2024年，国家数据局联合多部门先后发布两批共48个“数据要素×”典型案例，推动发挥数据在解决行业发展难题、促进行业效益提升等方面的重要作用，释放数据要素价值。目前，数据要素已在各行业开展了众多应用，如金融行业基于企业用电数据，经过脱敏、深度分析，掌握企业用电行为、用电缴费、用电水平、用电趋势等特征内容，为银行在信贷反欺诈、辅助授信、贷后预警等方面提供决策参考。数据要素已经成为赋能各行业数字化转型和智能化升级的重要基础。通过测算数据开发对经济增长的贡献度，可以看出，2022年第一、二、三产中数据对经济贡献度分别为0.32%、0.65%、1.69%，较2021年分别增长0.25%、0.49%、0.62%²，数据对经济增长的驱动效应逐步体现并加强。

（二）我国数据存储领域呈现快速发展态势

伴随人工智能、大数据等技术的高速发展，数据量呈指数级增长，海量数据对存储提出了更高的要求。我国加快发展存储产业，特别是先进存储技术成为重要发展方向，各地纷纷加快部署先进存储设施以提升数据存储和处理能力。

从存储政策方面看，我国各级政府积极制定出台产业发展政策。国家层面，工业和信息化部等6部门于2023年联合印发的《算力基

²数据来源：2024年中国信通院数字经济与工业经济领域深度观察

基础设施高质量发展行动计划》提出，到2025年，存储总量超过1800EB，先进存储容量占比达到30%以上。截至2023年底，我国存力规模达到约1.2ZB³，其中，先进存储容量占比超过25%。未来，我国存储容量将继续快速增长。2023年12月，国家发展改革委等部门发布的《关于深入实施“东数西算”工程加快构建全国一体化算力网的实施意见》提出，加快全闪存储等先进技术部署应用，推动西部地区承接全国范围的存储备份、离线分析等业务。地方层面，深圳市工业和信息化局发布《深圳市算力基础设施高质量发展行动计划(2024-2025)》，提出大力发展先进存储技术，并鼓励存算并举，规划建设与计算相匹配的存储体系。《上海市智能算力基础设施高质量发展“算力浦江”智算行动实施方案(2024-2025年)》提出，以“先进介质、高效架构、兼容生态、安全可信”为支撑，在智算中心内加快部署下一代先进存储技术。山东、广西、宁夏、湖南等地也提出了先进存储建设目标，以提升数据存储和处理能力，满足日益增长的存储和处理需求。

从存储产业方面看，我国存储产业规模较大。存储产业上游包含存储芯片与元器件，例如NAND Flash（SSD颗粒）、DRAM（动态随机存取存储器）等，中游包含存储整机和存储系统，下游是应用与服务。据IDC、Gartner等第三方机构预测，到2025年，我国存储产业规模将超万亿元，上游产业链超过2600亿元，中下游超过8000亿元。我国存储产业规模庞大，具有广阔的发展空间。

从存储技术方面看，我国数据存储介质研发、存储整机及系统产业领域已达到较高水平。存储介质领域，我国的NAND Flash和DRAM芯片已跻身全球主流水平。存储整机领域，国内存储厂商百花齐放，在全闪存储、分布式存储、备份存储、存储内生安全、统一文件系统、

³数据来源：数字中国发展报告（2023年）

数据编织、数据湖等存储技术领域已经具备国际竞争力。未来，存储技术将继续向“大容量、高性能、高效架构、安全可靠、绿色低碳”方向发展。

当然，我国存储领域也面临发展难题。主要包括新型人工智能存储技术研发有待进一步加强、存储与其他数据基础设施协同不够、先进存储设施建设步伐较慢等问题。例如，根据 IDC 数据，当前美国先进闪存在存储市场中的占比为 53%，相比之下，我国的同类产品占比为 18.2%，是全球平均水平的一半。因此，应发挥政策引导作用，加快我国存储产业发展，为网络强国、数据强国建设提供强有力支撑。

（三）新技术新模式新基建促进数据要素新发展

数据要素与行业新技术（以人工智能为代表）、新模式（以数据资产抵押贷款、信托等数据要素价值化新模式为代表）、新基建（以数据基础设施为代表）具有密切关系，这些新技术、新模式、新基建共同促进数据要素领域的新发展。

数据要素与人工智能之间联系紧密、相互促进。一是数据要素是人工智能发展的基础。例如 GPT 系列大模型具有大规模参数和强大的计算能力，其训练和推理过程需要海量的数据进行支撑。这些数据涵盖了文本、图像、语音等多种模态，为模型提供了丰富的学习素材。并且，数据的质量影响人工智能的智商。高质量的数据对于大模型的性能至关重要，数据的准确性、完整性和多样性直接影响模型的训练效果和泛化能力。二是人工智能推动数据要素的应用与挖掘。人工智能大模型通过深度学习技术，能够从海量数据中自动抽取和展现知识。这种能力使得数据要素中的隐藏价值得以被挖掘和利用，为各行业的智能化转型提供有力支持。大模型能够处理多种模态的数据，这种多模态融合的能力使得数据要素在跨领域、跨场景的应用中更加灵活和

高效。三是**数据要素与人工智能协同发展**。数据要素和人工智能大模型在发展中相互促进。一方面，高质量的数据推动了大模型技术的不断进步；另一方面，大模型技术的成熟又进一步提升了数据要素的应用价值。随着技术的不断发展，数据要素与大模型之间的融合创新将成为新的趋势。例如，在智能制造、智慧医疗等领域，数据要素与大模型的结合催生出一系列新的应用场景和商业模式。

数据要素价值化过程中形成了数据资产抵押贷款、数据信托等新模式。一是**数据要素的价值化包括资源化、资产化、资本化三个阶段**。数据要素经过数据采集、处理、分析和应用，形成具有经济价值和社会价值的**数据资源**，进而推动数据资产化、资本化进程。二是**数据要素促进金融市场发展**。依托数据资源形成的数据抵押贷款、数据信托等新模式，不仅丰富了金融产品的种类，也提高了金融服务的效率。三是**数据抵押贷款、数据信托等金融产品拓展了数据要素的应用场景**。通过创新数据资产的利用方式和融资渠道，可以提高数据要素的价值转化效率，为数据要素的应用和市场化配置提供更加便捷和高效的途径。

数据要素与数据基础设施如影随形、相生相伴、相辅相成、相互影响。一是**数据基础设施是支撑数据收集、存储、处理、分析和传输的物理和逻辑设施的总和**。它包括了网络设施、算力设施、存储设施、数据采集设施、数据流通设施、数据安全设施等，是数据要素得以产生、流动、整合和应用的必要环境。二是**数据要素与数据基础设施相互成就**。没有数据基础设施的支撑，数据要素将无法被承载和利用。反过来，没有数据要素的产生、流动和应用，数据基础设施也就没有存在的意义。三是**数据要素与数据基础设施也存在矛盾和盾的关系**。随着整个社会数字化水平不断提升，数据采集需求大量增加，数据量呈

指数级增长，这对网络带宽、算力、存储、流通、安全等数据基础设施建设提出更高要求。数据基础设施的不断升级又会促进数据要素更大规模地产生、流动和存储。例如，**数据和存储设施的关系就像水和瓶子一样**，随着水流量的大规模增加，需要更大容量的瓶子。这个瓶子的瓶口不能存在瓶颈，需要确保水能够快速流入流出。“瓶能装水亦能覆水”，并且“覆水难收”，因此，装水的瓶子需要更加坚固，使用这个瓶子也需要尽可能实现节能环保。综上，装水（数据）的瓶子（存储设施）需要具备大容量、高效率、安全可靠和绿色低碳等特点。

二、数据要素与人工智能相互促进，先进存储奠定坚实基础

数据是人工智能发展的三大核心要素之一，并已成为影响人工智能大模型效果的关键因素。人工智能大模型快速迭代创新，对数据的实时性、存储周期等也提出更高更复杂的要求，也促进了数据要素应用效能释放。

（一）人工智能加速数据要素升温，冷数据向温热数据转变

人工智能大模型推动冷数据变温变热。随着大模型参数规模越来越巨量化，训练所需数据资源体量也急剧增长。以 OpenAI 的 GPT 系列模型为例，GPT-1 数据集约为 4.8GB，GPT-2 数据集约为 40GB，而 GPT-3 数据集规模已超过 500GB。据相关研究⁴表明，互联网上可用的高质量文本数据或将在 2028 年耗尽，这一现象被称为“数据墙”，可能成为减缓人工智能进展的最大问题。因此，利用企业中积淀的丰富冷数据来训练模型，并从中挖掘新价值的机会正逐渐显现。虽然目

⁴来源：《经济学人》杂志文章《AI 公司很快将耗尽大部分互联网数据》（AI firms will soon exhaust most of the internet's data）

前尚处于起步阶段，但近期的技术突破正激励着企业以更加开放和创新的思维去探索冷数据在人工智能应用中的全新场景，促使冷数据变温变热。

人工智能助力各行业加速生产温热数据。人工智能的快速发展正围绕数据领域开辟全新机遇，而数据的焦点也多聚焦于热数据。当前，在人工智能带动下，更多热数据因人工智能而生。例如，智能电表数据，以前每小时抄一次就足够，但随着人工智能技术的进步，有必要提高数据采集频次，采集的范围也可以扩展到环境温度、湿度等数据，可用于预判用电负荷变化情况，优化电力调度。人工智能大模型的应用每天都在生成海量热数据，其中部分数据可直接用于合成优质数据集，继而再反馈喂给其他大模型用于训练、微调。

人工智能带来数据要素的发展也推动存储介质向全闪存化演进。人工智能大模型激发产生更大体量的温热数据，高频调用的需求暴涨，使得存储介质加快从机械硬盘（HDD）向固态硬盘（SSD）转变。随着机械硬盘向全闪存升级，存储介质的读写性能提升了百倍。全闪存存储是一种基于闪存技术的存储解决方案，它使用固态闪存器作为主要的数据存储介质，而不是传统的机械硬盘驱动器，提供了更高的数据访问速度和响应性，具备强大的技术优势。全闪存存储主打低延迟特性，对于高频交易和实时市场数据分析至关重要。目前，金融、电信、医疗和高性能计算等对数据访问速度要求极高的行业，是全闪存存储市场的主要关注者。因此，众多存储厂商均持续进行全闪存存储领域的技术创新和产品布局。

（二）人工智能提升数据存储价值，数据保存周期显著加长

人工智能大模型训练数据源深度持续拓展。在大模型的开发遭遇

能源瓶颈后，各大型企业开始从训练数据上寻找新的开发突破口，主要途径是扩大训练数据量和优化训练数据集，因此，人工智能大模型的训练数据集在规模和质量上逐渐提升。整体来看，人工智能大模型能力变强，可以训练更广维度与更长跨度的数据，极大拉长了数据存储周期，数据的长期价值被真正关注到。例如，气象数据已从存储 10 年，变成存储 30 年。

人工智能和数据要素发展推动存储技术向稳定可靠持久化方向演进。闪存存储、机械硬盘存储、磁带存储、蓝光存储均在向着大容量、高密度、低成本以及高可靠方向发展，为影像数据、文化遗产资料、科研数据等需要长期保存的数据提供支撑。同时，新型磁电介质存储、DNA 存储等作为一种前沿的存储技术，具有极高的数据密度和长期稳定性，正在逐渐走向成熟。

（三）高质量数据集助力人工智能，先进存储加速数据归集

人工智能大模型的训练和推理对高质量数据集的需求越来越大。高质量数据集是人工智能模型训练、推理和验证的关键基础，是按照特定标准，经过采集、清洗、归类和标注等智能化处理，并具备更新和维护机制的数据集合。谷歌研究发现，数据集大小至少与模型大小一样重要，数据集和模型大小应该约为 1:1 以达到给定数量的训练计算下的最佳性能。因此，未来人工智能的突破将得益于高质量数据。这是大模型性能提升、行业应用落地的关键，能否掌控相应的高质量数据集，直接决定能否构建产业竞争优势。

各行业高质量数据集支撑人工智能快速发展，先进存储加快数据集汇聚。例如，2024 年 3 月，北京国际大数据交易所牵头正式发布首批 100 个人工智能大模型高质量训练数据集，覆盖科技创新、金融服

务、医疗健康、医药研发、自动驾驶、气象服务、商业航天、影视创作等 20 多个应用场景，共计 100 余个语料数据，数据总量逾 150PB，涵盖了专业知识问答、古今文化书籍、互联网舆情资讯、多语种音视频、教学资源题库、科研数据、高清图片等多领域、多模态的数据语料。数据集的汇聚过程中，往往面临数据格式和访问协议多样的问题。为了让这些数据可以高效集中，需要数据存储设备支持多种不同的数据格式和访问协议，并且提供高写入带宽性能以实现这些多源异构的数据可以快速汇集在一起。因此，理想的存储硬件应支持多协议互通，确保统一的底层数据可以被不同协议/接口访问，避免因协议转换导致大量数据复制。先进存储设施满足这一要求，为人工智能训练提供进阶的数据管理能力，加速高质量数据集归集，从可视、可管、可用三个维度，帮助数据的拥有者和管理者以更加高效的方式来发挥数据价值。

（四）人工智能对存储提出高要求，新型人工智能存储取得突破

人工智能大模型的研发生产各阶段都对存储提出了新的要求。**数据采集环节**，由于原始训练数据规模巨大，且来源多样，需要构建大容量、低成本、高可靠的数据存储底座；**数据清洗阶段**，一个典型的百 TB 级大模型数据集，预处理时长超过 10 天，占比人工智能数据挖掘全流程的 30%，这就要求存储能够提供多协议、高性能的支持，用标准文件的方式完成海量数据的清洗和转换，以缩短数据预处理的时长；**模型训练环节**，主流预训练模型训练过程不稳定，易中断返工，需要用 Checkpoint（检查点）机制来确保训练回退到还原点，因此，存储能否快速地读写 Checkpoint 文件，也成了能否高效利用算力资源、提高训练效率的关键；**模型推理应用阶段**，存储需要提供比较丰

富的数据审核能力，以满足鉴黄鉴暴安全合规的诉求，保证大模型生成的内容是合法、合规的。

先进存储技术鼎力支撑人工智能发展。全闪存存储将带来性能大幅提升，加快人工智能大模型开发落地的速度。全闪存存储技术的高性能和大容量能够大幅提高数据读取速度，缩短模型训练和推理的时间。同时，全闪存存储技术的低延迟和高可靠性能够确保数据的安全性和完整性，减少数据丢失或损坏对模型性能的影响。数据编织、向量存储与近存计算等新兴数据处理技术，将最大程度降低企业整合数据、使用数据的门槛，满足资源的高效利用，降低行业接入人工智能大模型的难度。此外，先进存储内生安全体系将保护企业核心私密数据资产，让企业更加放心地使用人工智能大模型。

国内新型人工智能存储取得新突破，加速人工智能大模型迭代效率。随着人工智能大模型不断发展，数据存储技术也迎来了显著的创新和统一化趋势。各大存储厂商针对人工智能大模型生产的全流程，进行了有针对性的产品设计优化。综合各大厂商的存储技术创新，可以发现技术方向较为统一，都是基于人工智能大模型训练推理的全流程，在高性能、弹性扩展、长记忆、知识库、高可靠等方向进行了全新的设计创新。国内存储领军企业已经具备了人工智能领域的较强国际竞争力，据全球权威的人工智能基准测试 MLPerf™ 发布的存储测试结果显示，国内新型人工智能存储产品获得全球第一。

三、数据资产化推动数据要素价值释放，先进存储成为数据资产保险库

数据已成为数字经济时代最重要的资产之一。我国率先开展数据资产评估入表与登记，率先开展基于数据资产的抵押融资，为数据要素价值化创造了基础条件。正如金融资产存在银行保险箱里一样，数

据资产也需要有安全可靠的保险库。先进存储设施正是存储数据资产的保险库，是存储数据资产的最重要的基础设施，为数据的资产化、产权化提供坚强保障。

（一）数据资产评估入表和产权登记取得重要突破

从全球看，大多数国家尚未出台数据资产管理规定，部分国际组织开始初步探索。2024年7月，由我国主导，联合多国共同制定的国际标准化组织（ISO）数据资产领域首个国际标准《资产管理 数据资产管理指南》（ISO 55013:2024 Asset management — Guidance on the management of data assets）正式发布。负责修订国际财务报告准则（IFRS）的国际会计准则理事会（IASB）正逐步探索如何将无形资产尤其是数据资产，纳入财务报告体系。例如，通过发布讨论稿或征求意见稿，探讨数据资产的识别、计量、列报和披露等问题。这种探索体现了国际社会对数据资产经济价值的广泛认可。

近年来，我国加快推动数据资产管理评估工作，在全球各国中率先出台企业数据资产入表政策文件，为其他国家提供了有益参考。2023年至今，我国财政部门发布有关数据资产的4份政策文件，即3份有关数据资产的规定，并指导中国资产评估协会制定印发1份指导意见。2023年8月，《财政部关于印发〈企业数据资源相关会计处理暂行规定〉的通知》（财会〔2023〕11号）正式发布，自2024年1月1日起施行。在数据资产评估方面，2023年9月，《中评协关于印发〈数据资产评估指导意见〉的通知》（中评协〔2023〕17号）发布，明确了数据资产价值的评估方法包括收益法、成本法和市场法三种基本方法及其衍生方法。2024年1月，《财政部关于印发〈关于加强数据资产管理的指导意见〉的通知》（财资〔2023〕141号）正式发布。鼓励公共服务机构将依法合规持有或控制的、具有资产属性的公共数据

资源纳入资产管理范畴。次月,《关于加强行政事业单位数据资产管理的通知》(财资〔2024〕1号)发布,专门针对行政事业单位数据资产管理进行规定,明晰管理责任,规范管理行为,实现从企业数据资产到行政事业单位数据资产管理全覆盖。

我国数据资产入表规定生效后,企业积极开展数据资产评估入表实践。2024年一季度,在我国境内上市的超过5300家上市公司中,有25家公司的财务报表中首次披露了数据资产“入表”情况。但是,后续有7家发布公告称因填报错误等问题更正了一季报并剔除了数据资源相关数据。剩下的18家上市公司在一季度财报中披露的数据资源入表合计金额1.03亿元,其中无形资产项下0.79亿元,开发支出项下0.18亿元,其他部分列入存货项下。2024年8月9日,中国移动在半年报中首次披露了数据资产入表情况,是第一家实施数据资产入表工作的央企,数据资产入表金额为7000万元,其中无形资产2900万元,开发支出4100万元。半年报指出包括数据资源在内的无形资产,以成本计量。数据资产预计使用寿命与软件、著作权一样,均为2至5年。8月15日,中国联通半年报发布,其中,数据资源科目金额为8476.39万元,全部列入开发支出中。中国联通在半年报指出:“2024年6月30日,本集团开发支出余额中包含尚在开发中的数据资源约人民币0.85亿元,主要包含为现有数据产品和服务提供支撑的行业数据库和模型等。”8月20日,中国电信披露半年报,其中开发支出科目明细显示数据资产入表金额为1.05亿元,成为基础电信企业中入表金额最多的,标志着中国电信在数据要素价值释放和数字化转型方面迈出了重要一步。至此,三家基础电信企业已经全部披露了数据资产入表情况。根据万得公司(Wind)数据,截至2024年8月31日,我国境内上市公司共有39家(不含事后更正的5家)

在半年报中披露了数据资产入表情况，入表金额合计 12.04 亿元，环比一季度增长约 11 倍。

各地区探索依托数据资产开展融资。2023 年以来，数据资产融资的案例不断出现。例如，2023 年 12 月，在中国人民银行盐城市分行指导下，江苏银行盐城分行与盐城市大数据集团举行全市首笔数据使用权抵押贷款发放暨战略合作协议签约仪式，成功发放 1000 万元贷款支持盐城市数据产业业务发展。2024 年 6 月，神州数码公司成功将金服云数据产品作为数据资产，纳入企业财务报表并获得建设银行深圳分行授信融资 3000 万元。这是全国大中型数据资产质押融资的代表性案例，已通过中国人民银行动产融资统一登记公示系统完成数据资产质押登记。

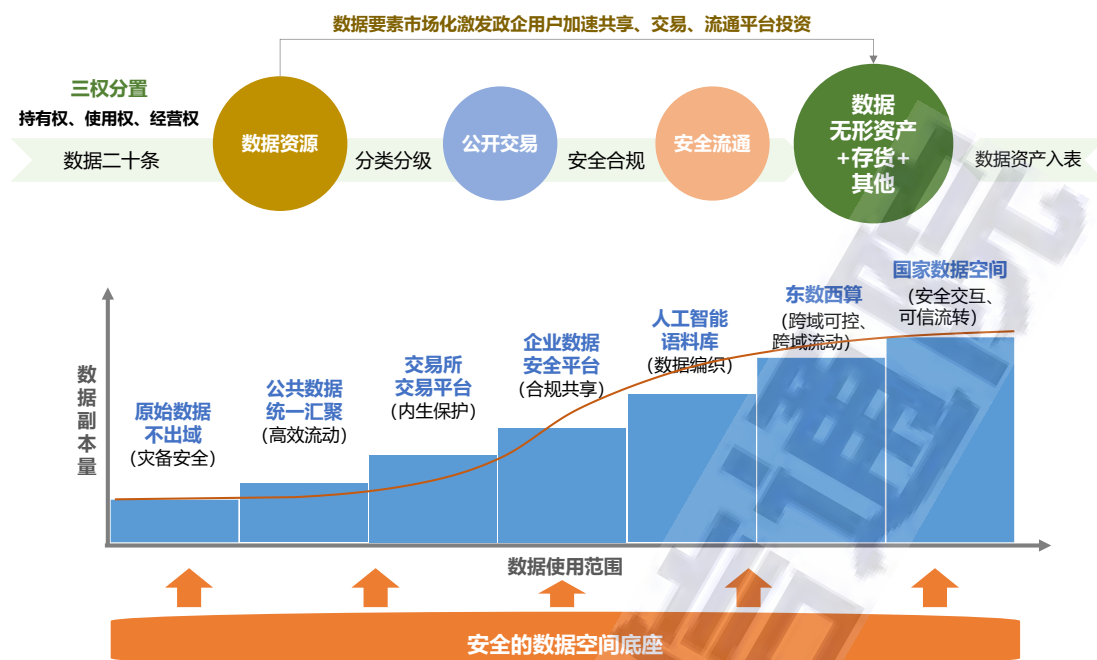
另外，我国数据知识产权方面取得较大进展。在国家知识产权局组织下，2022 年 12 月，北京、上海、江苏、浙江、福建、山东、广东、深圳 8 个省市开启数据知识产权试点工作，上线数据知识产权登记平台。2023 年 12 月，按照《国家知识产权局办公室关于深化数据知识产权地方试点工作的通知》(国知办函规字[2023]999 号)要求，在 8 个已有试点地方的基础上，新增天津市、河北省、山西省、安徽省、河南省、湖北省、湖南省、贵州省、陕西省等 9 个地方共同作为数据知识产权试点地方。截至 2024 年 7 月，全国 17 个开展数据知识产权试点工作的省市，共颁发数据知识产权登记证书超过 8700 件，数据知识产权质押融资金额超过 55 亿元⁵。2024 年 9 月，国家知识产权局制数据知识产权登记证书正式开放申请，这标志着国家版数据知识产权登记证书正式推出，全国数据知识产权登记迈出重要一步。同时，数据知识产权登记证书在司法判决中发挥了重要作用。数据知识

⁵数据来源：国务院新闻办公室“推动高质量发展”系列主题新闻发布会，2024 年 7 月 29 日

产权登记证书可用于维护证书持有人的权益。例如，数据堂（北京）科技股份有限公司与隐木（上海）科技有限公司间不正当竞争纠纷一案先后经过北京互联网法院一审以及北京知识产权法院二审。涉案的数据堂公司数据取得了北京知识产权保护中心发放的《数据知识产权登记证书》，北京知识产权法院认为，在没有相反证据证明的情况下，该证书初步证明了数据堂公司为数据的合法持有人，数据堂公司有权就登记的数据进行权益主张。最终，数据堂公司在该案二审中胜诉。

（二）数据资产流通交易带来数据副本量快速增加

在原始数据不出域条件下，数据副本并未增加。但是，当数据一旦开始进入共享、开放、交易等流通环节则必将带来数据副本量的快速增加。特别是，数据作为资产进行流通也将带来数据副本量的进一步膨胀。发挥数据资产价值，要求数据副本尽可能多的流通。存储设施作为数据资产“持有权”的实际载体，可以作为数据资产的实体标的物，在数据要素市场中发挥数据资产兑现的价值，包括：创建受数据持有者和交易条款控制的可信数据空间，受持有方隐私、合规策略控制的数据交易节点，成为数据交易市场的副本流转节点。



来源：中国信息通信研究院

图 1 数据使用范围与对应的数据副本量增加情况

（三）先进存储为数据资产提供安全可靠的保险库

数据是重要的资产，不仅是决策的依据，也是支撑技术创新和服务的基础，应把数据资产存放在先进存储设施这一保险库里，确保数据资产安全。同时，先进存储设施也为数据产权保护奠定了坚实基础。

先进存储设施为数据资产提供安全可靠、可兼容、可扩展和高性价比的承载基础。一是先进存储设施具有很强的可靠性。随着数据存储技术的不断发展，闪存等先进存储介质正在逐步替代传统的机械硬盘，这些新型介质具有更高的存储密度和更好的耐用性，具有很强的可靠性，为数据资产提供了更加安全的存储介质。二是先进存储设施具有良好的兼容性。先进存储设施支持多协议存储，能够兼容不同来源和类型的数据，为数据的多样化存储提供了可能。这有助于汇聚多样化的数据资源，形成规模化的数据资产，从而更好地挖掘数据价值。三是先进存储设施具有很高的可扩展性。随着数据量的不断增加，企业对存储容量的需求也在不断增长。先进存储设施可扩展性强，能够

根据客户的实际需求灵活调整存储容量，确保数据资产能够得到充分的存储和利用。四是先进存储设施具有很好的性价比。先进存储设施具有容量大、性价比高等特点，客户通过采用先进存储解决方案，可以大幅降低数据存储和管理的成本，同时享受高效、便捷的数据服务。

先进存储设施为数据产权保护提供基础性支撑。首先，通过建立数据基础设施中的防勒索机制，健全数据存储的容灾备份功能，利用先进存储技术打造数据流通基础设施的坚强底座，从而增强对数据产权的保护能力。其次，先进存储设施支持数据的分类存储和权限管理，能够根据数据的敏感性和重要性设置不同的访问权限，确保只有合法用户才能访问其有权访问的数据。这有助于防止数据泄露和滥用，保护数据产权拥有者的相关权益。再次，先进存储设施提供审计和日志记录功能，记录所有用户的访问和操作日志。这有助于事后审计和监控用户的访问行为，确保数据访问的合规性和安全性，保护数据产权安全。

四、先进存储设施是数据基础设施的关键一环

数据从产生到利用涉及到数据的采集、存储、处理、交易、流通等各个环节，数据基础设施建设支撑着数据要素的全生命周期管理。其中，先进存储设施是数据基础设施的重要组成部分，为促进数据要素的开发利用、推动数字经济高质量发展提供了重要基础。

（一）数据基础设施是关键性载体，存储是“六个底座”之一

数据基础设施是数据“供得出、流得动、用得好、保安全”的重要保障。数据基础设施是从数据要素价值释放的角度出发，在网络、算力、存储等设施的支持下，面向社会提供一体化数据采集、汇聚、处理、流通、应用、运营、安全保障服务的一类新型基础设施。通过

构建完善的数据基础设施，可以提高数据的采集效率和质量，实现不同区域、不同领域数据资源的整合与共享，推动数据在各行业、各机构间的合规流通，促进数据资源的有效利用，为数据要素市场化配置奠定重要基础。数据基础设施的完善和发展正在推动数据服务千行百业、深度融入社会生产生活，为经济社会的数字化转型和高质量发展提供有力支撑。

数据基础设施的建设为数据的全生命周期管理构筑了坚实支撑。其中，**网络设施**是数据流通的基础，通过光纤、5G、卫星互联网等先进技术，为数据提供高速、稳定、广泛的连接能力。**算力设施**是数据处理的核心，包括通用算力、智能算力和超级算力等多种类型的计算技术和资源，为数据提供高效、敏捷的处理能力。**存储设施**是数据保存和管理的核心。它利用先进的存储技术和设备，为数据提供稳定、可靠、高效的存储空间。**数据采集设施**通过传感器、RFID、图像采集等技术将物理世界的信息转化为数字信号，为数据流通提供了原始的数据源。**数据流通设施**是实现数据共享和流通的关键，通过利用数据空间、区块链、隐私计算等技术，打通数据共享流通的堵点，实现数据在不同主体间的“可用不可见”、“可控可计量”。**数据安全设施**是保护数据安全的重要手段，它采用隐私保护、数据加密、数字身份认证等技术手段，为数据提供全方位的安全保障。

存储设施与网络、算力、数据采集、数据流通、数据安全等设施共同构成数据基础设施的“六个底座”。没有存储设施，其他数据基础设施就成了“空中楼阁”。**网络设施**是数据流通的桥梁，它连接着存储设施、算力设施以及其他数据基础设施，存储设施的性能直接影响网络数据存取响应速度，决定着数据流通的效率。**算力设施**利用强大的计算能力对存储在存储设施中的数据进行高效处理和分析，数据

存储是算力资源发挥作用的前提和基础，算力的发挥离不开包括内存和外置存储在内的存储设施。数据采集设施负责数据采集和初步处理，而数据存储设施为数据采集设施提供了存储保障。数据流通设施包括数据开放、共享、交易等平台，存储设施为平台中的数据提供基础的支撑，确保数据资源流动的畅通性。数据安全设施是数据保护盾，存储设施可为海量数据提供安全、可靠的存储空间，为数据提供安全之基。



来源：中国信息通信研究院

图 2 数据基础设施组成部分

（二）数据基础设施建设提速提质，先进存储占比不断提升

网络设施进入提速升级新阶段。我国 5G 网络建设深入推进，核心网能力持续增强，已建成全球规模最大、技术领先的 5G 网络。根据《数字中国发展报告（2023 年）》数据，截至 2023 年底，我国 5G 基站数达 337.7 万个，同比增长 46.1%；平均每万人拥有 5G 基站 24 个，较上年末提高 7.6 个百分点。光纤宽带网络方面，我国已建成全球规模最大的光纤宽带网络，为数据传输提供了高速、稳定的网络环

境。其中具备千兆网络服务能力的 10G PON 端口数达 2302 万个，增幅达 51.2%，已形成覆盖超 5 亿户家庭的能力。1000Mbps 及以上接入速率用户达 1.63 亿户，在固定宽带接入用户占比达到 25.7%，较上年末提高 10.1 个百分点。

算力设施布局持续优化。随着“东数西算”工程进入到全面建设阶段，各类算力加快向国家枢纽节点集聚，算力资源利用效率大幅提升。截至 2023 年底，全国在用数据中心标准机架超过 810 万架，算力总规模达到 230EFLOPS，居全球第二位，算力总规模近 5 年年均增速近 30%⁶。算力供给结构逐步优化，通用算力、智能算力、超级算力等多元算力资源融合发展，其中智能算力规模占比持续提升，截至 2023 年底，智能算力规模达到 70EFLOPS，增速超过 70%。超级计算机性能大幅提升，全国累计建成国家级超算中心 14 个⁷，科学计算和工程计算能力显著提升，有力保障了数据高性能计算资源的需求。

⁶数据来源：数字中国发展报告（2023 年）

⁷数据来源：数字中国发展报告（2023 年）

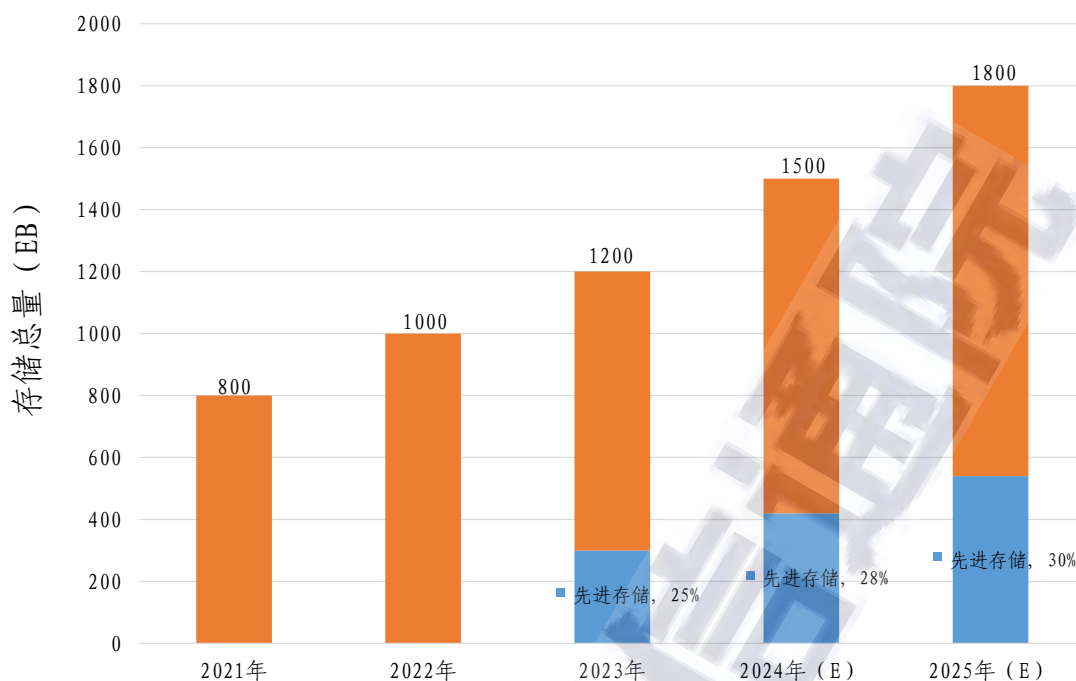


来源：工业和信息化部

图 3 2019 年-2023 年我国数据中心机架数

存储设施需求持续增长,先进存储占比不断提升。随着人工智能、大数据等技术的飞速发展,海量数据的存储和处理需求不断增长,对存储设施提出了更高的要求。先进存储设施具有大容量、高效率、高可靠和绿色低碳特点。地方纷纷加快部署先进存储设施以提升数据存储和处理能力。在政策指引下,我国存储设施正呈现出与网络设施、算力设施等一体化发展之势。截至 2023 年底,我国存力规模达到约 1.2ZB⁸,其中,先进存储容量占比超过 25%。根据《算力基础设施高质量发展行动计划》发展指标,到 2024 年底和 2025 年底,我国数据存储总容量将分别达到 1500EB 和 1800EB,其中先进存储占比分别达到 28%和 30%。

⁸数据来源:数字中国发展报告(2023年)



来源：工业和信息化部、中国信息通信研究院

图 4 2021 年-2025 年我国存储总量及先进存储占比

（三）存力中心建设成为典型模式，各地建设取得重要成效

随着数字化转型的纵深推进，数据的集中管理和高效流通变得至关重要，在此背景下，构建以先进存储设施为基石的数据基础设施体系是数字化转型的必然要求，能够有效促进数据的横向流通与纵向沉淀，激活数据的潜在价值，为我国数字经济发展提供重要支撑。其中，存力中心成为各地先进存储设施落地的典型模式。

存力中心是集数据归集、保护、治理、加工与供给于一体的综合平台，为各种应用场景提供高效的数据存储和管理服务。通过聚合场景与数据资源，存力中心逐渐成为“数据要素汇聚开发中心、价值中心、产业聚能中心、数据资产中心、数据安全保障中心”集合体，高效管理数据要素在“采、存、算、管、用、产”各个环节，为算力中心、人工智能大模型及数据交易提供高质量数据，充分激活海量数据价值，释放算力潜能。

当前各地区加快采用先进的存储技术和架构部署建设先进存力中心，以提供高性能、高可靠、高扩展的数据存储解决方案。其中，“存算一体化”已成为存力中心建设的主流模式，该模式通过深度融合数据存储与计算处理功能，实现数据的即时分析和决策支持，提高了数据处理的效率和响应速度。存算一体化模式通过容灾、备份、非结构化数据管理等手段，实现数据中心从“重建设”到“重运营”的转变，从单纯的“存数”到“算数”和“用数”转变，有效打通数据融合、构建产业生态。

专栏：存力中心建设典型实践

全国一体化算力网络国家（贵州）主枢纽中心数算一体存力中心：2023年，贵安新区大数据科创城产业集群有限公司建设了全国一体化算力网络国家（贵州）主枢纽中心数算一体存力中心，将产生的海量数据进行实时地收集与保存，并为算力中心和数据交易所持续提供高质量数据供给，充分激活数据价值，释放算力潜能，最终实现覆盖数据要素“采、存、算、管、用、产”各个环节的全生命周期管理。此外，存力中心提供自主可控、全面防勒索、存储安全、双重保护、数据加密、可靠的灾难恢复、网络安全和架构安全八大安全保障，不仅保障了数据的安全和可靠性，也为数据的高效利用和价值实现提供了坚实的支撑。

同时，凭借在区位、能源、成本及数据安全等方面的独特优势，贵阳贵安吸引海量东部及异地数据资源汇聚到贵州先进存力中心，并吸引数据治理、数据开发、数据资产运营等全产业链企业落地贵州，共同推进贵阳贵安数算一体化产业集群的建设。该项目的实施不仅巩固了贵安新区作为全国一体化算力网络关键节点的地位，还将建立一个服务全国、辐射全球的数据服务平台，为贵州数字经济

的高质量发展注入了强大动力。

西部（重庆）科学城先进数据中心存储一体化存力中心：重庆市政府和璧山区人民政府共同建立西部科学城先进数据中心，该数据中心采用的是一体化存力方案。一体化存力方案可以总结为六大特性：场景融合一体化、数据联邦一体化、生命周期一体化、数据服务一体化、数据安全一体化、数据绿色一体化。

存储一体化存力方案可为用户带来多场景的存储支撑，支持冷热分级存储跨区域无感调度。用户可通过可视化界面，进行从数据产生到归档的全生命周期管理。此外，将液冷技术与存储设施相结合，可为数据中心构建存算一栈式液冷解决方案，以满足绿色低碳需求。

五、发展建议

（一）推进高质量数据集建设，强化新型人工智能存储创新应用

发挥人工智能技术优势，对大量文本、图像、音频等数据进行自动标注和分类，批量构建高质量数据集。加强各方协作，探索面向人工智能大模型的先进存储技术产品标准建设。以市场侧实际需求为牵引，加快推进先进存储在更多人工智能应用场景中落地，引导制造、能源、交通等领域的大型企业加快打造一批“经济效益好、社会影响大、示范效应强”的新型人工智能存储应用。

（二）完善数据资产评估体系，增加数据资产的存储安全性评估

推动企事业和行政单位数据资产化，在数据资产评估中纳入数据资产安全可靠存储、风险管理等指标。优先推动企业开展数据资产入

表与登记,遵循《企业数据资源相关会计处理暂行规定》等政策要求,确保数据资产入表的合规性和准确性。逐步推动行政事业单位将公共数据资产纳入资产管理范畴,建立健全公共数据资产入表与登记的相关制度。优化数据资产评估体系,强化针对数据资产存储安全性、可靠性的考量,加强数据风险管理,防止数据泄露或被篡改。

(三) 加快建设先进存储设施, 优化我国数据基础设施建设布局

进一步优化完善我国数据基础设施建设布局,将存储基础设施纳入我国数据基础设施的关键组成部分,明确其在数据存储、处理、分析及应用中的关键作用,促进存储与算力等其他数据基础设施协同发展。制定国家层面的存储基础设施建设规划,确保存储基础设施建设的系统性和前瞻性。根据区域经济发展需求、产业布局及数据流量特征,合理规划并适度超前建设我国存储基础设施。在重点行业、关键领域和示范区部署先进存储基础设施应用示范项目,推动技术成果的转化和应用推广。强化存储基础设施的安全保障,加强存储基础设施的物理安全、网络安全和运维安全,提升设施的安全防护能力和应急响应能力。加强数据加密、访问控制、安全审计等安全措施,确保存储数据的安全性、完整性和可用性。

中国信息通信研究院 产业与规划研究所

地址：北京市海淀区花园北路 52 号

邮编：100191

电话：010-68021205

传真：010-68033959

网址：www.caict.ac.cn

