

国家数据基础设施建设指引

2024 年 12 月

前 言

党的十八大以来，以习近平同志为核心的党中央，敏锐把握新一轮科技革命和产业变革的新机遇，统揽中华民族伟大复兴的战略全局和世界百年未有之大变局，对发展数字经济作出重大部署，擘画了新时代数字中国建设的宏伟蓝图。

党的二十届三中全会明确提出“建设和运营国家数据基础设施，促进数据共享”。各地区各部门认真贯彻落实习近平总书记重要指示精神，积极探索数据基础设施建设，为数据要素市场化配置改革、建设全国一体化数据市场奠定了良好基础。同时也要看到，数字经济蓬勃发展对数据流通利用和价值释放提出了新的更高的要求，迫切需要更好发挥有为政府和有效市场作用，构建兼顾效率和公平、适应数据要素特征、发挥数据价值效用的国家数据基础设施。

按照党中央、国务院决策部署，国家发展改革委、国家数据局、工业和信息化部在充分调研的基础上，组织编制了《国家数据基础设施建设指引》，力争在当前情况下，说清楚数据基础设施的概念、发展愿景和建设目标，指导推进数据基础设施建设，推动形成横向联通、纵向贯通、协调有力的国家数据基础设施基本格局，打通数据流通动脉，畅通数据资源循环，促进数据应用开发，培育全国一体化数据市场，夯实数字经济发展基础，为数字中国建设提供有力支撑。

一、概念内涵

纵观人类经济发展史，每一轮产业变革都会孕育新的基础设施。农业经济时代，基础设施主要是农田水利设施。工业经济时代，公路、铁路、港口、机场、电力系统等成为关键基础设施。数字经济时代，网络设施、算力设施、应用设施等构建了数字基础设施。当前，数据成为关键生产要素，催生新的技术—经济范式，重塑产业发展方式，推动数字基础设施向数据基础设施延伸和拓展。建设和运营国家数据基础设施，进一步促进数据“供得出、流得动、用得好、保安全”，对于支撑数据基础制度落地、构建全国一体化数据市场、培育发展新质生产力具有重要意义。

国家数据基础设施是从数据要素价值释放的角度出发，面向社会提供数据采集、汇聚、传输、加工、流通、利用、运营、安全服务的一类新型基础设施，是集成硬件、软件、模型算法、标准规范、机制设计等在内的有机整体。国家数据基础设施在国家统筹下，由区域、行业、企业等各类数据基础设施共同构成。网络设施、算力设施与国家数据基础设施紧密相关，并通过迭代升级，不断支撑数据的流通和利用。

二、发展愿景

（一）主要目标

国家数据基础设施是数据基础制度和先进技术落地的重要载体。在数据流通利用方面，建成支持全国一体化数据市场、保障数据安全自由流动的流通利用设施，形成协同联动、规模流通、高效

利用、规范可信的数据流通利用公共服务体系。在算力底座方面，构建多元异构、高效调度、智能按需、绿色安全的高质量算力供给体系。在网络支撑方面，构建泛在灵活接入、高速可靠传输、动态弹性调度的数据高速传输网络。在安全方面，构建整体、动态、内生的安全防护体系。在应用方面，支持传统行业转型升级，赋能人工智能等新兴产业发展。总体实现“汇通海量数据、惠及千行百业、慧见数字未来”的美好愿景。

（二）推进路径

当前，我国数据基础设施处于起步建设阶段，围绕流通利用业务场景，各地方各行业各领域探索形成多种有针对性的技术方案和解决路径，并在不断迭代发展。在推动技术设施化过程中，要注重发挥有为政府和有效市场双重作用，坚持自上而下布局、自下而上探索双向协同，鼓励大胆创新，支持先行先试，加快技术收敛，推动技术规模化部署、系统化应用，为构建高速互联、高效调度、开放普惠、安全可靠的国家数据基础设施奠定坚实基础。

2024—2026年，利用2—3年左右时间，围绕重要行业领域和典型应用场景，开展数据基础设施技术路线试点试验，支持部分地方、行业、领域先行先试，丰富解决方案供给。制定统一目录标识、统一身份登记、统一接口要求的标准规范，夯实数据基础设施互联互通技术基础。完成国家数据基础设施建设顶层设计，明确国家数据基础设施建设的技术路线和实践路径。

2027—2028年，建成支撑数据规模化流通、互联互通的数据

基础设施，数网、数算相关设施充分融合，基本形成跨层级、跨地域、跨系统、跨部门、跨业务的规模化数据可信流通利用格局，实现全国大中型城市基本覆盖。

到 2029 年，基本建成国家数据基础设施主体结构，初步形成横向联通、纵向贯通、协调有力的国家数据基础设施基本格局，构建协同联动、规模流通、高效利用、规范可信的数据流通利用体系，协同构筑数据基础设施技术和产业良好生态，国家数据基础设施建设运营体制机制基本建立。

三、总体功能

数字中国、数字经济、数字社会建设提出了数据资源化、要素化、价值化要求，国家数据基础设施围绕打造可信流通、高效调度、高速互联、安全可靠的体系化能力，持续赋能各行业数据融合与智能化发展。

（一）数据可信流通：开放普惠的数据流通

国家数据基础设施需要打造低成本、高效率、可信赖的流通环境，便于人、物、平台、智能体等快速接入，在符合统一目录标识、统一身份登记、统一接口要求的基础上，实现数据在不同组织、行业之间安全有序流动，精准匹配数据供需关系。面向电子商务、金融支付、跨境物流、航运贸易、绿色低碳、气象服务等典型场景，创新融合数据应用，同时符合相关法律法规、社会伦理、个人隐私保护等要求。

（二）高效算力供给：多元异构的算力协同

算力资源多元异构、异地分布、动态变化，给大规模计算任务的统一调度与任务协同带来挑战。面向“东数西算”等场景中对异构异构异地算力的调度需求，需要建立多元异构算力统筹调度的能力，促进算力和运力的高度融合，推进算力资源之间的无缝对接与协同计算，提高整体计算效率与资源利用率，实现算力最优配置与动态调整。

（三）数据高速传输：高效弹性的数据传输网络

高效弹性的传输网络可为数字金融、智慧医疗、交通物流、大模型训练和推理等核心场景数据传输流动提供高速稳定服务。国家数据基础设施在高效弹性传输网络的支撑下，能够显著提升数据交换性能，降低数据传输成本，为数据大规模共享流通提供高质量通道。

（四）全程安全可靠：动态全面的安全保障

数据采集、汇聚、传输、加工、流通、利用、运营等多样化活动，涉及多方主体、多个环节，需要在开放环境下对数据进行整体、动态保护。国家数据基础设施需要构建标准化、多层次、全方位的安全防护框架，推动安全防护由静态保护向动态保护、由边界安全向内生安全、由封闭环境保护向开放环境保护转变，形成贯穿数据全生命周期各环节的动态安全防护能力，系统保障数据基础设施相关的网络、算力、数据、应用安全。

四、总体架构

（一）技术架构

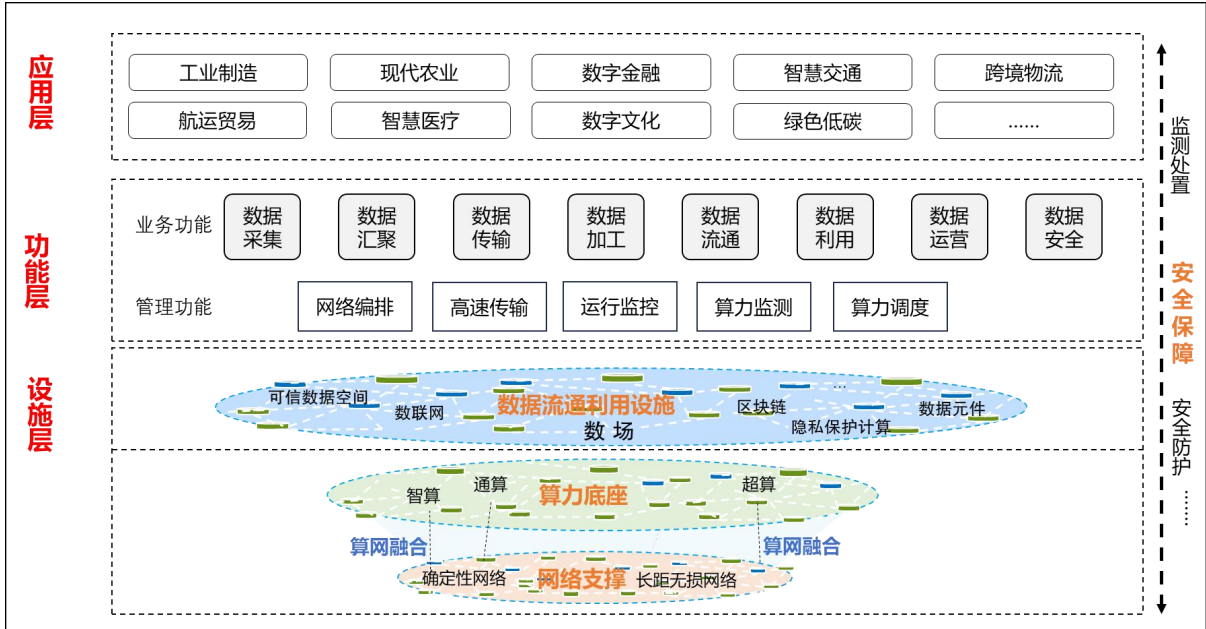


图 1 数据基础设施及网络、算力设施总体架构图

国家数据基础设施具有数据采集、汇聚、传输、加工、流通、利用、运营、安全八大能力。在数据采集方面，支持通过传感器、业务系统等手段采集相关数据。在数据汇聚方面，通过标识编码解析、数据目录等，对数据进行高效接入、合理编目，实现数据广泛汇聚、存储和发布。在数据传输方面，支持节点即时组网、数据高效传输。在数据加工方面，为参与方提供高效便捷、安全可靠的数据清洗、计算服务，建立数据质量控制和评估能力，提高数据处理环节效率。在数据流通方面，通过数据分类分级策略实现共享、交易等流通功能，为不同行业、不同地区、不同机构提供可信流通环境。在数据利用方面，为数据应用方提供数据分析、数据可视化等能力，进一步降低数据应用门槛。在数据运营方面，提供数据登记、

监督管理、数据认证、合规保障等功能，有效支撑全国一体化数据市场有序运行。在数据安全方面，提供动态全过程数据安全服务，包括防窃取、防泄露、防滥用、防破坏等。在赋能方面，促进数据多场景应用、跨主体复用，赋能工业制造、现代农业、跨境数字货币、数字金融、智慧医疗、智慧交通、跨境物流、航运贸易、绿色低碳等行业领域。

其中，数据流通利用设施是国家数据基础设施的重要组成部分，为跨层级、跨地域、跨系统、跨部门、跨业务数据流通利用提供安全可信环境，包括可信数据空间、数场、数据元件、数联网、区块链网络、隐私保护计算平台等技术设施。网络设施、算力设施适应数据价值释放需要，向数据高速传输、算力高效供给方向升级发展。安全保障体系是国家数据基础设施安全可靠运行的保障，包括监测预警、信息通报、应急处置等相关制度、能力和队伍建设。

（二）主要构成

国家数据基础设施以行业、区域数据基础设施为主体，以企业数据基础设施为重要组成。企业数据基础设施是指服务企业生产、运营、管理的数据平台，包括采集、存储、处理、管理等相关硬件和软件系统，以及企业整合、协同关联数据方形成的数据服务平台。行业数据基础设施是指覆盖某一行业领域，服务行业内企业、用户及利益相关者，实现数据资源化、要素化、价值化的各类设施，包括行业数据流通交易平台、行业数据归集平台、行业数据公共服务平台等。区域数据基础设施是指覆盖本地区，服务区域内企业、用

户及利益相关者，实现数据资源化、要素化、价值化的各类设施，包括数据归集平台、数据资源管理服务平台、公共数据开放平台、公共数据运营平台等。国家在企业、行业、区域数据基础设施的基础上，组织建设基于统一目录标识、统一身份登记、统一接口要求的数据流通利用设施底座，搭建数据流通利用设施管理平台，以及建设数据产权登记、公共数据运营、数据资源管理、数据流通交易、算力资源监测调度等基础公共服务的平台。这些设施相互贯通、协同推进，共同促进国家数据基础设施建设发展。

五、重点方向

（一）建设数据流通利用设施底座

按照统一目录标识、统一身份登记、统一接口要求，建设数据流通利用设施底座。建立覆盖政府、行业、企业等主体及国家、省、市、县等层级的全国一体化的分布式数据目录，形成全国数据“一本账”，支撑跨层级、跨地域、跨系统、跨部门、跨业务的数据有序流通和共享应用。建立全国一体化的分布式数字身份体系，规范身份标识生成、身份注册和认证机制。建立统一的数据资产凭证、交易凭证结构、生成与验证机制，支持利用区块链、加密技术、智能合约等手段提高凭证的可溯性和信任性。构建标准化、规范化的交互接口，实现数据基础设施的互联互通。建设数据泛在接入体系，支持数据资源、参与主体、第三方服务更大规模接入。建立与 IPv6 等网络标识兼容的数据标识体系。建立数据目录分类分级管理机制，加强数据分类管理和分级保护。

（二）建设数据高效供给体系

在数据标注产业的生态构建、能力提升和场景应用等方面先行先试。链接公共数据，主动公开的企业数据和个人数据，以及各类高质量数据集，对社会形成统一的数据资源开放目录。研究制定高质量数据集建设相关标准，从数据生成、注释定义到数据管理的全过程，确保数据标注的准确性和数据模型的专业性。制定高质量数据标注与交付规则，提高训练数据质量。支持农业、工业、交通、金融、自然资源、卫生健康、教育、科技、民航、气象等行业领域打造高质量数据集。因地制宜推进公共数据运营平台集约化、标准化建设，推进公共数据的规模化、常态化供给。推进数据资源管理服务平台互联互通，完善平台标准，促进平台间互操作，实现全国数据资源的跨领域、跨层级、跨区域流通利用。构建集成数据采集、存储、清洗、标注、管理、应用等功能的一体化数据基础通用工具平台，提升数据加工效率，保证数据质量。支持各地积极建设政务服务大模型，推动政务服务智能化。

（三）建设数据可信流通体系

建立高效便利可信的数据流通机制，促进数据大规模、低成本、安全自由流通。支持建设企业可信数据空间、行业可信数据空间、城市可信数据空间，探索建设个人可信数据空间、跨境可信数据空间。支持基础好、有条件、意愿强的行业和城市，先行先试数场建设。鼓励行业、地方积极探索建设区块链网络、隐私保护计算平台等新技术设施。支持因地制宜，探索数联网、数据元件等数据流通

利用设施建设。支持建设数据流通交易公共服务平台。支持探索建设数据跨境流动基础设施。建立数据流通准入标准规则，鼓励探索数据流通安全保障技术、标准、方案。

（四）建设数据便捷交付体系

加强数据交易场所体系设计，统筹数据交易场所优化布局。支持数据交易场所创新发展，鼓励各类数据进场交易。构建集约、高效的数据交付基础设施，为场内集中交易和场外分散交易提供低成本、高效率、可信赖的数据交付环境。促进各类数据交易所、交易平台互联互通。推动数据价值贡献度评估、数据集推荐匹配、数据产品差异性分析等技术创新，实现供需精准匹配和便捷交付。鼓励各地提升数据加工、测试、建模验证、安全实验等社会化服务能力，打造产学研用“一公里”工作圈。

（五）建设行业数据应用体系

加强场景牵引，建设面向工业制造、现代农业、数字金融、智慧医疗、智慧交通、跨境物流、航运贸易、卫生健康、绿色低碳、气象服务、数字文化等重点行业领域的数字应用体系，发挥企业主体作用，促进行业数据应用创新。培育基于数据要素的新产品和服务，促进数据多场景应用、跨主体复用，实现知识扩散、价值倍增。

六、算力底座

（一）推进算力资源科学布局

加快推动通用算力、智能算力、超级算力等多元异构算力的绿

色发展、有机协同。促进各类新增算力向国家枢纽节点集聚，强化枢纽节点国家算力高地定位。建设全国一体化算力网监测调度平台。探索采用存算分离架构建设新型智算中心和新材料大数据中心。

（二）推进东中西部算力协同

加强新兴网络技术创新应用，优化网络计费方式，降低东西部数据传输成本，促进东部中高时延业务向西部转移。推进算力互联互通，构建算力多级调度策略引擎，实现跨平台、跨层级、跨区域的算力资源混合部署和统一调度，促进算力资源高效对接，提升数据汇聚、处理、流通、交易效率。推动国家枢纽节点和需求地之间400G/800G高带宽全光连接，引导电信运营商等提升“公共传输通道”效能，推进算网深度融合。

（三）推进算力与数据、算法融合创新

不断壮大数算产业生态体系，助力打造具有国际竞争力的数字产业集群。推动行业数据和算力协同，实现数据可信流通，提升数据处理能力和治理水平。建立健全算法开发利用机制，积极开展大模型创新算法及关键技术研究，提升数据分析能力，降低大模型计算的算力消耗水平。

（四）推进算力与绿色电力融合

推动实现“瓦特”产业向“比特”产业转化，强化枢纽节点与非枢纽节点的协同联动，支持绿电资源丰富的非枢纽节点融入全国一体化算力网建设。加强大型风光基地和算力枢纽节点协同联动，把绿色电力转换成绿色算力。积极推进风光绿电资源消纳，助力实

现碳达峰碳中和。支持利用“源网荷储”等新型电力系统模式。加强数据中心智慧能源管理,开展数据中心用能监测分析与负荷预测,优化数据中心电力系统整体运行效率。探索绿电直供新模式,有序开展绿电、绿证交易。

(五) 推进算力发展与安全保障协同

推动建设国家算力网基础安全保障平台,打造一体化的安全保障服务能力。打造网络和数据安全攻防演习靶场,推动国家枢纽节点地区定期开展网络和数据安全攻防演习。建设算力网安全应用技术试验场。强化国家枢纽节点自主防护能力,统一应急处置、统一安全监测、统一运行监控,构筑全生命周期的安全管控措施。

七、网络支撑

建设高速数据传输网,实现不同终端、平台、专网之间的数据高效弹性传输和互联互通,解决数据传输能力不足、成本较高、难以互联等问题。支持电信运营商、相关科研机构、国家大科学装置等,叠加虚拟化组网、网络协议创新和智能化任务调度等云网融合技术,形成多方快速组网和数据交换能力,支持面向数据传输任务的弹性带宽和多量纲计费。

推动传统网络设施优化升级,有序推进 5G 网络向 5G-A 升级演进,全面推进 6G 网络技术研发创新。在东中西部地区均衡布局国际通信出入口局,加快扩展国际海缆、陆缆信息通道方向。建设时延确定、带宽稳定保障、传输质量可靠的确定性网络。布局“天地一体”的卫星互联网。

八、安全防护

国家数据基础设施安全保障体系建设重点是构建多层次、全方位、立体化的国家数据基础设施安全保障框架，贯穿数据生命周期全流程，帮助各参与方提升数据安全保障能力，确保数据的可信性、完整性和安全性。

在国家数据基础设施安全保障层面，实现可信接入、安全互联、跨域管控和全栈防护等安全管理，建立网络安全风险和威胁的动态发现、实时告警、全面分析、协同处置、跨域追溯和态势掌控能力，提供应对芯片、软件、硬件、协议等内置后门、漏洞安全威胁的内生防护能力。加强对合作伙伴、运维人员、平台用户等数据安全内部风险的防范应对。加强对入侵渗透、拒绝服务、数据窃取、勒索投毒等外部威胁的应急响应。

在数据流通利用安全层面，综合利用隐私保护计算、区块链、数据使用控制等技术手段，保证数据的可信采集、加密传输、可靠存储、受控交换共享、销毁确认及存证溯源等，规避数据隐私泄露、违规滥用等风险。加强算法、模型、数据的安全审计，增强模型鲁棒性和安全性，保证高价值、高敏感数据“可用不可见”“可控可计量”“可溯可审计”，确保贯穿数据全生命周期各环节安全。

九、组织保障

（一）健全政策保障体系

建立健全数据基础制度体系，加快出台数据产权、流通交易、收益分配、安全治理等政策文件。在新型基础设施规划安排下，研

究制定国家数据基础设施建设规划。加大中央投资对国家数据基础设施建设的支持力度。各地区、各部门要在数据基础设施规划布局、资金安排、课题研究方面给予重点支持。积极引导社会资本力量参与国家数据基础设施建设。

（二）加快技术创新探索

支持有条件的行业和地区开展先行先试，探索建设数据基础设施。鼓励企业和科研机构加大研发投入，加快数据流通利用关键技术攻关和重大成果转化。通过国家重点研发项目课题立项、揭榜挂帅、数据技术创新大赛等方式推动技术创新。

（三）强化标准和人才支撑

强化标准支撑，研究制定数据基础设施相关标准规范。鼓励企业、社会团体、科研机构参与数据基础设施国际标准的制定工作。加强与 ISO、IEC、ITU、IEEE、3GPP 等国际标准化组织的合作，推动数据领域高水平专家在国际组织任职。推动人才队伍建设，建立数据人才评价标准和评选机制。

附录：技术术语解释

（一）数据流通利用技术

在数据流通利用领域，目前常用的技术路线主要包括隐私保护计算、区块链、数据使用控制等。

1.隐私保护计算

隐私保护计算指在保证数据提供方不泄露原始数据的前提下，对数据进行分析计算的一类信息技术，保障数据在产生、存储、计算、应用、销毁等数据流转全过程的各个环节中“可用不可见”。隐私保护计算的常用技术方案有安全多方计算、联邦学习、可信执行环境、密态计算等；常用的底层技术有混淆电路、不经意传输、秘密分享、同态加密等。

2.区块链

区块链是分布式网络、加密技术、智能合约等多种技术集成的新型数据库软件，具有多中心化、共识可信、不可篡改、可追溯等特性，主要用于解决数据流通过程中的信任和安全问题。

3.数据使用控制

数据使用控制是指在数据的传输、存储、使用和销毁环节采用技术手段进行控制，如通过智能合约技术，将数据权益主体的数据使用控制意愿转化为可机读处理的智能合约条款，解决数据可控的前置性问题，实现对数据资产使用的时间、地点、主体、行为和客体等因素的控制。

（二）数据流通利用实践方案

在数据流通利用领域，目前业界的实践方案主要包括可信数据空间、数场、数联网、数据元件等。

1.可信数据空间

可信数据空间是基于共识规则，联接多方主体，实现数据资源共享共用的一种数据流通利用设施，是数据要素价值共创的应用生态，是支撑构建全国一体化数据市场的重要载体。可信数据空间须具备数据可信管控、资源交互、价值共创三类核心能力。

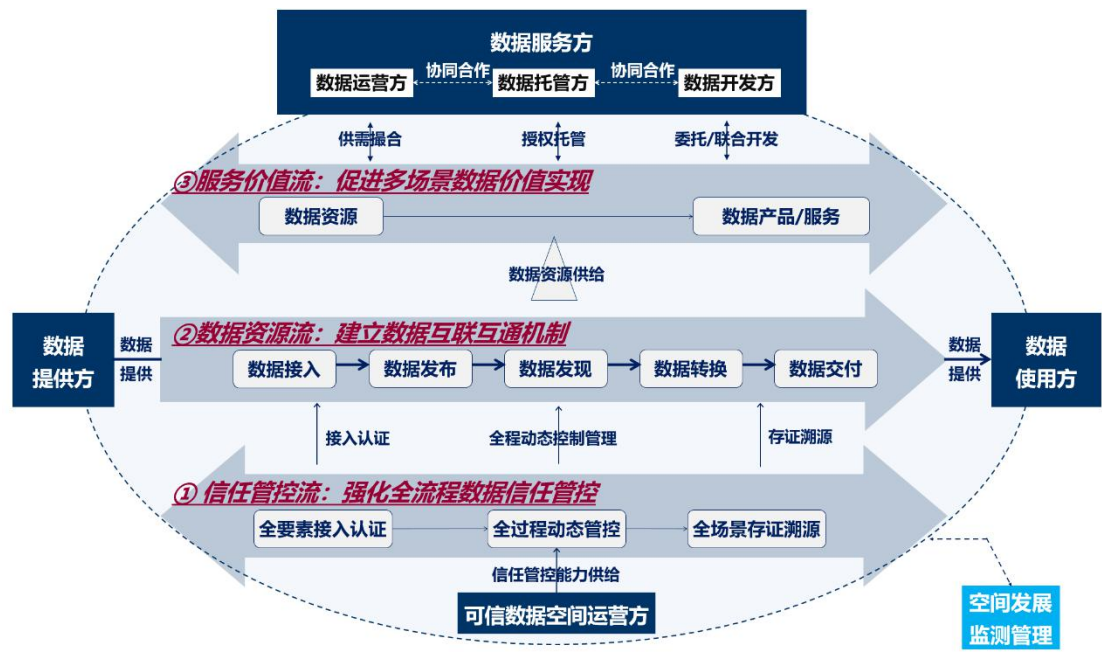


图 2 可信数据空间架构图

2.数场

数场是依托开放性网络、算力和隐私保护计算、区块链等各类关联功能设施，面向数据要素提供线上线下资源登记、供需匹配、交易流通、开发利用、存证溯源等功能，支持多场景应用的一种综合性数据流通利用设施。数场以高效流通、价值释放、繁荣生态为

目标，具备实现数据可见、可达、可用、可控、可追溯等能力，具有开放性、融合性、扩展性等特点。

数场从点、线、面、场、安全五个维度构建标准化技术框架。点是数据主体进入数场的接入点。线是数场内连接各主体、各平台的高速数据传输网，实现数场内各主体之间的互联互通。面是数场中数据主体、传输网络的集合，是实现数据大规模流通、高效安全利用的核心。由点到线、由线到面构成数场基础设施。场是基于数场基础设施构建的数据应用、场景化创新，以及相关能力、流程、规范的统称。安全是覆盖点、线、面、场的动态全流程保护措施。数场在技术架构上包括接入点、功能平台、管理平台、安全保障、网络传输等基础服务平台。

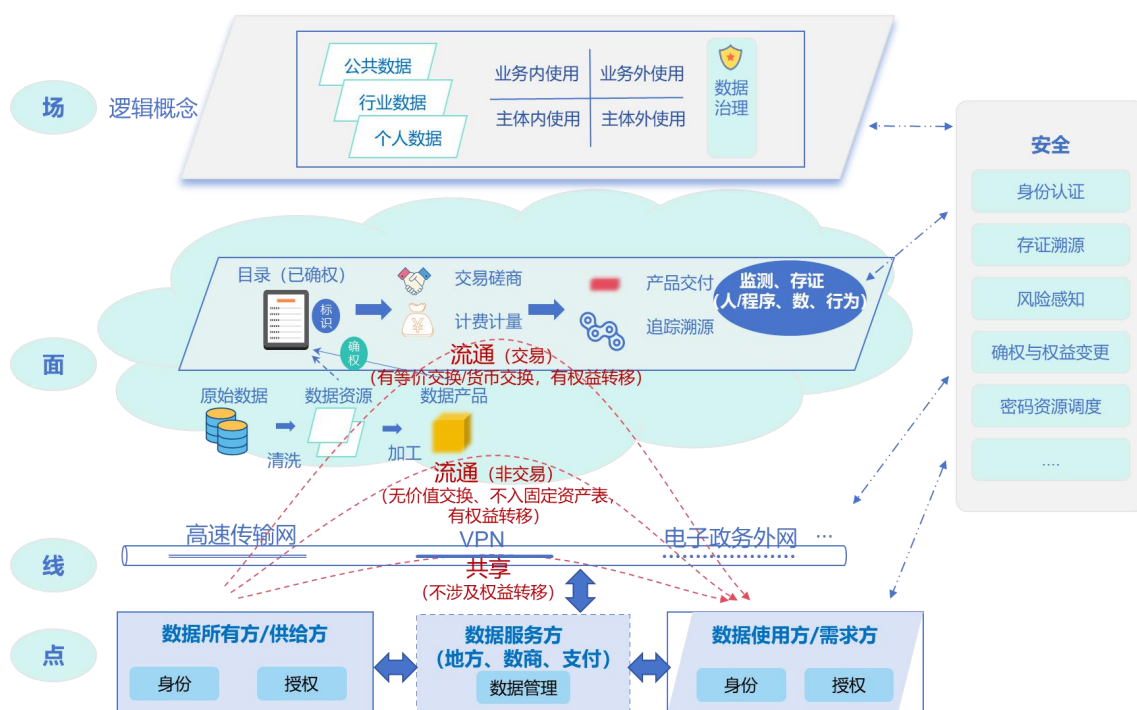


图3 数场功能架构图

3.数联网

数联网由数据流通接入终端、数据流通网络、数据流通服务平台构成，提供一点接入、广泛连接、标准交付、安全可信、合规监管、开放兼容的数据流通服务。

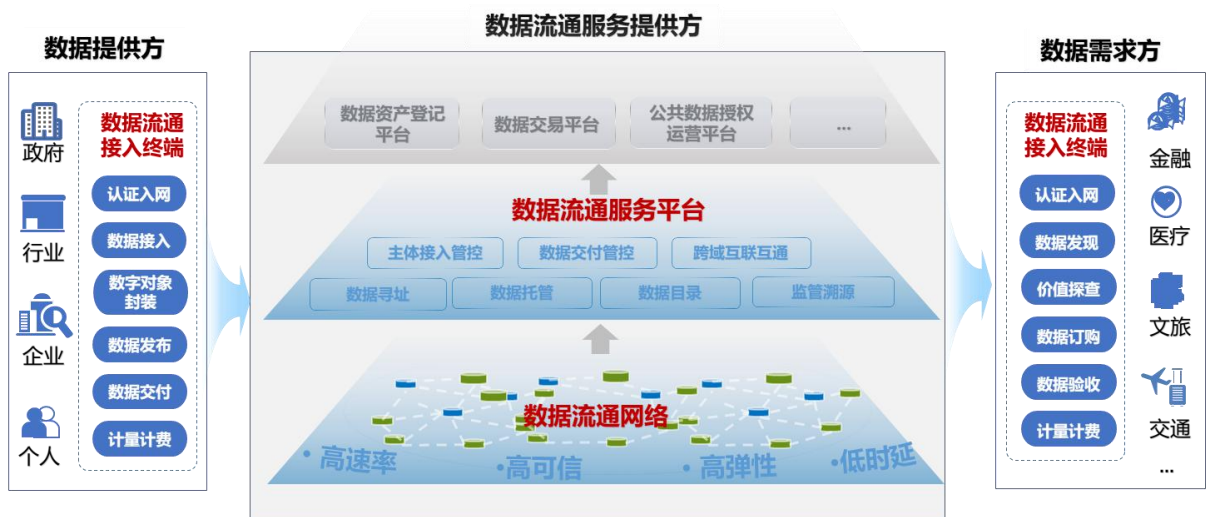


图 4 数联网功能架构图

4.数据元件

数据元件提供统一标准、自主可控、安全可靠、全程监管的数据存储和加工服务，支持采用标准化工序完成数据产品规模化加工、生产和再利用，适用于大规模数据加工和生产场景。数据元件作为连接数据供需两端的“中间态”，将原始数据与数据应用“解耦”，基于数据元件相关组件，实现从数据归集到数据元件加工交易全生命周期的数据要素开发和管控。

（三）数据安全技术

数据安全技术为数据收集、存储、处理、传输、共享和销毁等全生命周期提供安全保障，包括数据备份与恢复、应用数据加密、

数据泄露检测、流转监测、身份认证与访问控制、数据脱敏、数据水印、数据安全态势感知等。