Translation

In the following document, three Chinese government ministries describe an ambitious plan to make all or most of China's public data, data disclosed by corporations, and personal data discoverable and usable in a single catalog by 2029. The effort involves creating a standard data identifier system, identity authentication scheme, and interface, as well as a secure network infrastructure for accessing the data. The document does not, however, identify which ministries will be responsible for implementing and enforcing the plan, and explains that the scheme is still in the exploratory phase and that actual construction of the catalog will not begin until 2027.

Title
Guidelines for National Data Infrastructure Construction
国家数据基础设施建设指引

Authors
Chinese National Development and Reform Commission (NDRC; 国家发展和改革委员会; 国家发展改革委; 国家发改委), National Data Administration (国家数据局), and Ministry of Industry and Information Technology (MIIT; 工业和信息化部; 工信部)

Source
NDRC website. The Guidelines are dated December 31, 2024 and were released to the public on January 6, 2025.

The Chinese source text is available online at:
https://www.ndrc.gov.cn/xxgk/zcfb/tz/202501/P020250106319424108511.pdf
An archived version of the Chinese source text is available online at: https://perma.cc/KX7B-KUQS

Translation Date	**Translator**	**Editor**
April 22, 2025	Etcetera Language Group, Inc.	Ben Murphy, CSET Translation Manager

Guidelines for National Data Infrastructure Construction

December 2024

Preface

Since the 18th Party Congress [in 2012], the Party Central Committee, with Comrade Xi Jinping as the core, has astutely grasped the new opportunities presented by the new round of scientific and technological (S&T) revolution and industrial transformation, taken into account the overall strategic situation of the great rejuvenation of the Chinese nation and the great changes in the world unseen in a century (世界百年未有之大变局), made major arrangements for the development of the digital economy, and outlined a grand blueprint for the construction of a Digital China in the new era.

The Third Plenum of the 20th Chinese Communist Party (CCP) Central Committee clearly proposed that we "construct and put into operation national data infrastructure to promote data sharing." All regions and departments have conscientiously implemented the spirit of the important instructions of General Secretary Xi Jinping and actively explored the construction of data infrastructure, laying a good foundation for the reform of marketized allocation of data factors of production (要素) and the construction of a national integrated data market. At the same time, we must also see that the vigorous development of the digital economy has put forward new and higher requirements for the circulation, utilization, and unleashing of the value of data. It is urgently necessary to allow assertive government and effective markets (有为政府和有效市场) to better play their roles and build a national data infrastructure that takes into account efficiency and fairness, adapts to the characteristics of data factors of production, and gives full play to the value and effectiveness of data.

In accordance with the decisions and arrangements of the Party Central Committee and the State Council, the National Development and Reform Commission (NDRC), the National Data Administration, and the Ministry of Industry and Information Technology (MIIT) have organized the compilation of the *Guidelines for National Data Infrastructure Construction* based on extensive research. This document strives to clarify the concepts, development vision, and construction goals of data infrastructure under the current circumstances, guide and promote data infrastructure construction, and promote the formation of a basic national data infrastructure pattern that features horizontal connection, vertical integration, and strong coordination. It strives to open up the arteries of data circulation, smooth the circulation of data resources, promote data application development, cultivate a national integrated data market, lay a solid foundation for the development of the digital economy, and provide strong support for the construction of a Digital China.

Contents

1. Conceptual implications

Throughout the history of human economic development, every round of industrial transformation will give birth to new infrastructure. In the agricultural economy era, infrastructure mainly consisted of farmland and water conservancy facilities. In the industrial economy era, roads, railways, ports, airports, power systems, and other facilities became key infrastructure. In the digital economy era, network facilities, computing power (compute) facilities, application facilities, and other facilities built the digital infrastructure. At present, data have become a key factor of production, giving birth to a new techno-economic paradigm, reshaping the way industries develop, and promoting the extension and expansion of digital infrastructure to data infrastructure. The construction and operation of national data infrastructure will further promote the supply, flow, utilization, and security (“供得出、流得动、用得好、保安全”) of data, which is of great significance for supporting the implementation of basic data systems, building a national integrated data market, and cultivating and developing new productive forces (新质生产力).

National data infrastructure is a new type of infrastructure that provides data collection, aggregation, transmission, processing, circulation, utilization, operation, and security services to society with the aim of unleashing the value of data factors of production. It is an organic whole that integrates hardware, software, model algorithms, standards and specifications, and mechanism design. National data infrastructure is coordinated by the state and is composed of various types of data infrastructure such as regional, industrial, and enterprise infrastructure. Network facilities and compute facilities are closely related to national data infrastructure, and through iterative upgrades, they continuously support the circulation and utilization of data.

2. Development vision

(1) Main objectives

National data infrastructure is an important vehicle for the implementation of basic data systems and advanced technologies. In terms of data circulation and utilization, we will build circulation and utilization facilities that support the national integrated data market and ensure the secure and free flow of data, forming a public service system for data circulation and utilization that is coordinated, interconnected, standardized, reliable, and features large-scale circulation and efficient utilization. In terms of compute base, we will build a high-quality compute supply system that is diverse and heterogeneous, efficiently scheduled, intelligent, on-demand, green, and secure. In terms of network support, we will build a high-speed data transmission

network with ubiquitous and flexible access, high speed, and reliable transmission, and dynamic and elastic scheduling. In terms of security, we will build a holistic, dynamic, and endogenous security protection system. In terms of applications, we will provide support for the transformation and upgrading of traditional industries and enable the development of emerging industries such as artificial intelligence (AI). The overall goal is to achieve the beautiful vision of "connecting massive data, benefiting thousands of industries, and gaining insight into the digital future."

(2) Path of advancement

At present, China's data infrastructure is in the stage of initial construction. Focusing on the circulation and utilization business scenarios, various localities, industries, and fields have explored and formed a variety of targeted technical solutions and approaches and are constantly iterating and developing them. In the process of promoting the building of technological infrastructure, we must focus on giving full play to the complementary functions of assertive government and effective markets, adhere to the two-way coordination of top-down layout and bottom-up exploration, encourage bold innovation, support pioneering trials, accelerate technological convergence, and promote large-scale deployment and systematic application of technology. This is done in order to lay a solid foundation for building national data infrastructure that is high-speed, interconnected, efficient, open, inclusive, secure, and reliable.

From 2024 to 2026, we will use about 2–3 years to carry out pilot tests on different technical approaches to data infrastructure, with a focus on important industries and fields and typical application scenarios. We will provide support for certain regions, industries, and fields to carry out trials first and enrich the supply of solutions. We will formulate standards and specifications for unified catalog identifiers (目录标识), unified identity registration, and unified interface requirements to lay a solid technical foundation for the interconnection and interoperability of data infrastructure. We will complete the top-level design of national data infrastructure construction and clarify the technical approaches and practical paths of national data infrastructure construction.

From 2027 to 2028, we will build data infrastructure supporting large-scale data circulation and interconnection and fully integrate data network- and data compute-related facilities. We will basically form a pattern of trusted circulation and utilization of data at scale across levels, regions, systems, departments, and lines of business, achieving basic coverage of large and mid-sized cities across the country.

By 2029, we will have basically established the main structure of the national data infrastructure and have initially formed the basic pattern of national data

infrastructure, featuring horizontal connectivity, vertical integration, and strong coordination. We will have built a data circulation and utilization system featuring coordinated links, large-scale circulation, efficient utilization, and standardized and reliable operations and have jointly built a favorable ecosystem for data infrastructure technology and the data infrastructure industry. We will have basically established national data infrastructure construction and operation institutions and mechanisms.

3. Overall functions

The construction of Digital China, a digital economy, and a digital society has put forward requirements for making data into a resource, making it into a factor of production, and giving it value (数据资源化、要素化、价值化). National data infrastructure focuses on building systematic capabilities for trusted circulation, efficient scheduling, high-speed interconnection, and security and reliability as well as the continuous empowerment of data integration and intelligentized (智能化) development in various industries.

(1) Trusted circulation of data: Open and inclusive data circulation

The national data infrastructure needs to create a low-cost, high-efficiency, and reliable circulation environment to facilitate rapid access by people, things, platforms, and intelligent agents (智能体). On the basis of complying with the requirements of unified catalog identifiers, unified identity registration, and unified interface requirements, it must achieve the safe and orderly flow of data between different organizations and industries and accurately align data supply and demand. It must innovatively integrate data applications for typical scenarios such as e-commerce, financial payments, cross-border logistics, shipping and trade, green and low-carbon industries, and meteorological services, while complying with relevant laws and regulations, social ethics, and personal privacy protection requirements.

(2) Efficient compute supply: Coordination of diverse and heterogeneous compute

Diverse, heterogeneous, geographically dispersed, and dynamically changing compute resources pose challenges to the unified scheduling and task coordination of large-scale compute tasks. In order to meet the scheduling requirements for heterogeneous and remote compute in scenarios such as "East-West Compute Transfer,"[1] it is necessary to establish the ability to coordinate and dispatch multiple

[1] Translator's note: "East-West Compute Transfer" (“东数西算”; literally, "eastern data, western compute") refers to an initiative to build computing power in western China, where land and electricity are relatively cheap, in support of data centers located along China's densely populated and developed east coast.

heterogeneous compute resources, promote the high degree of integration of compute and transportation capacity, promote seamless connection and collaborative computing between compute resources, improve overall compute efficiency and resource utilization, and achieve optimal configuration and dynamic adjustment of compute.

(3) High-speed data transmission: Efficient and flexible data transmission network

An efficient and flexible transmission network can provide high-speed and stable services for data transmission and circulation in core scenarios such as digital finance, intelligent healthcare, transportation and logistics, and large model training and inference. Supported by an efficient and flexible transmission network, national data infrastructure can significantly improve data exchange performance, reduce data transmission costs, and provide a high-quality channel for large-scale data sharing and circulation.

(4) End-to-end security and reliability: Dynamic and comprehensive security assurance

Diverse activities such as data collection, aggregation, transmission, processing, circulation, utilization, and operations involve multiple parties and multiple stages. Therefore, they require holistic and dynamic protection of data in an open environment. The national data infrastructure must build a standardized, multi-level, and comprehensive security protection framework and promote the transformation of security protection from static protection to dynamic protection, from perimeter security to endogenous safety and security (内生安全), and from closed environment protection to open environment protection. The infrastructure must form dynamic security protection capabilities that run throughout all aspects of the data life cycle, and it must systematically ensure the security of networks, compute, data, and applications related to data infrastructure.

4. Overall architecture

(1) Technology architecture

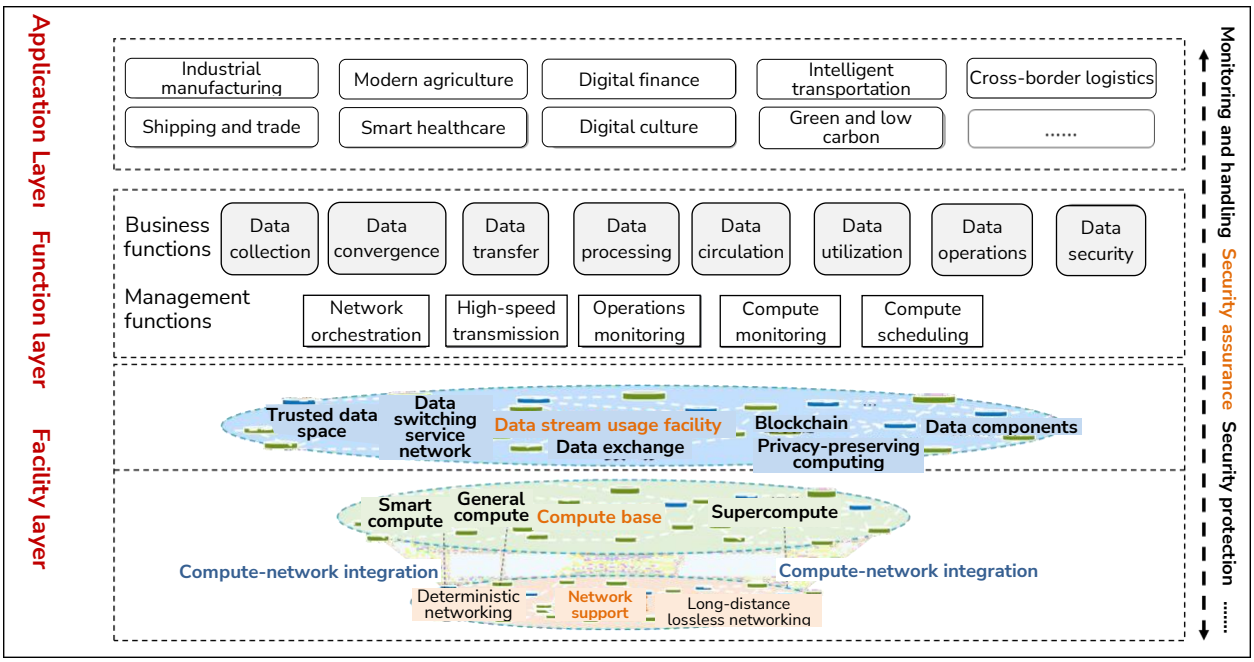

Figure 1: Overall architecture of data infrastructure, network, and compute facilities

National data infrastructure has eight major capabilities: data collection, convergence, transmission, processing, circulation, utilization, operation, and security. In terms of data collection, the infrastructure supports the collection of relevant data through sensors, business systems (业务系统), and other means. In terms of data aggregation, data can be efficiently accessed and reasonably cataloged through identifier coding analysis, data catalogs, and other means to achieve extensive data aggregation, storage, and release. In terms of data transmission, the infrastructure supports instant node networking (节点即时组网) and efficient data transmission. In terms of data processing, the infrastructure provides participants with efficient, convenient, secure, and reliable data cleaning and compute services, establishes data quality control and evaluation capabilities, and improves the efficiency of data processing. In terms of data circulation, the infrastructure uses data categorization and grading policies to achieve circulation functions such as sharing and trading, providing a trusted circulation environment for different industries, regions, and institutions. In terms of data utilization, the infrastructure provides data application users with data analysis, data visualization, and other capabilities to further lower the thresholds of data applications. In terms of data operations, the infrastructure provides functions such as data registration, supervision and management, data authentication, and compliance assurance, effectively supporting the orderly operation of the national integrated data market. In terms of data security, the infrastructure provides dynamic end-to-end data security services, including protection against theft, leaks, abuse, and destruction. In terms of empowerment, the infrastructure promotes the multi-scenario

application and cross-entity reuse of data, enabling industrial manufacturing, modern agriculture, cross-border digital currency, digital finance, intelligent healthcare, intelligent transportation, cross-border logistics, shipping and trade, the green and low-carbon industry, and other industries.

Data circulation and utilization facilities are an important part of the national data infrastructure. They provide safe and trusted environments for the circulation and utilization of data across levels, regions, systems, departments, and lines of business, including trusted data spaces, data exchanges (数场), data components, data switching service networks (DSSNs; 数联网), blockchain networks, privacy-preserving computing platforms, and other technical facilities. Network facilities and compute facilities adapt to the needs of unleashing the value of data, and upgrade and develop towards high-speed data transmission and efficient compute supply. The security assurance system is the guarantee for the secure and reliable operation of national data infrastructure. It includes relevant systems, capabilities, and teams for features such as monitoring and early warning, information notification, and emergency response.

(2) Main components

National data infrastructure is mainly composed of industry and regional data infrastructure, with enterprise data infrastructure as an important component. Enterprise data infrastructure refers to the data platforms that serve the production, operation, and management of an enterprise, including related hardware and software systems, such as those for collection, storage, processing, and management, as well as the data service platforms formed by the enterprise's integration of and collaboration between relevant data parties. Industry data infrastructure refers to various facilities that cover a certain industry or field, serve enterprises, users, and stakeholders within the industry, and make data into a resource, make it into a factor of production, and give it value, including industry data circulation and trading platforms, industry data collection platforms, and industry data public service platforms. Regional data infrastructure refers to various facilities that cover a region, serve enterprises, users, and stakeholders in the region, and make data into a resource, make it into a factor of production, and give it value, including data collection platforms, data resource management service platforms, public data open platforms, and public data operation platforms. Based on the data infrastructure of enterprises, industries, and regions, the state organizes the construction of a foundation of data circulation and utilization facilities based on unified catalog identifiers, unified identity registration, and unified interface requirements, builds a data circulation and utilization facility management platform, and builds platforms for basic public services such as data property registration, public data operations, data resource management, data circulation and trading, and compute resource monitoring and scheduling. These facilities are

advanced in an interconnected and coordinated manner to jointly promote the construction and development of national data infrastructure.

5. Key directions

(1) Build a foundation of data circulation and utilization facilities

We should build a foundation of data circulation and utilization facilities based on unified catalog identifiers, unified identity registration, and unified interface requirements. We should establish a nationwide integrated distributed data catalog covering government, industry, enterprises, and other entities as well as the national, provincial, city, and county levels, form a "one-ledger" national data catalog, and support the orderly flow and shared application of data across levels, regions, systems, departments, and lines of business. We should establish a nationwide integrated and distributed digital identity system and standardize identity generation, identity registration, and authentication mechanisms. We should establish unified data asset certificates, a trading certificate structure, and generation and verification mechanism and support the use of blockchain, encryption technology, smart contracts, and other means to improve the traceability and trustworthiness of certificates. We should build standardized and specification-based interactive interfaces to achieve the interconnection and interoperability of data infrastructure. We should build a ubiquitous data access system and support larger-scale access to data resources, participating entities, and third-party services. We should establish a data identifier system that is compatible with IPv6 and other network identifier systems. We should establish a data catalog-based categorized and graded management mechanism and strengthen categorized data management and graded protection.

(2) Build an efficient data supply system

We should take the lead in conducting pilot programs in areas such as ecosystem construction, capability improvement, and scenario-based applications in the data labeling industry. We should link public data, voluntarily disclosed corporate data, and personal data as well as various high-quality data sets and form a unified open catalog of data resources for society. We should research and formulate relevant standards for the construction of high-quality data sets and ensure the accuracy of data labeling and the specialization of data models throughout the entire process from data generation, annotation, and definition to data management. We should establish high-quality data labeling and delivery rules and improve the quality of training data. We should support the creation of high-quality data sets in industries and fields such as agriculture, industry, transportation, finance, natural resources, health, education, S&T, civil aviation, and meteorology. We should promote the intensive (集约化) and

standardized construction of public data operation platforms in accordance with local conditions and promote the large-scale and normalized supply of public data. We should promote the interconnection and interoperability of data resource management service platforms, improve platform standards, promote interoperability between platforms, and realize the cross-domain, cross-level, and cross-regional circulation and utilization of national data resources. We should build an integrated general purpose data infrastructure tools platform that integrates data collection, storage, cleaning, labeling, management, application, and other functions, improve data processing efficiency, and ensure data quality. We should support local governments in actively building large-scale government service models and promote intelligentized government services.

(3) Build a trusted data circulation system

We should establish an efficient, convenient, and trusted data circulation mechanism and promote the large-scale, low-cost, secure, and unrestricted (自由) circulation of data. We should support the construction of enterprise trusted data spaces, industry trusted data spaces, and city trusted data spaces and explore the construction of personal trusted data spaces and cross-border trusted data spaces. We should support industries and cities with good foundations, preconditions, and strong will in constructing data exchange pilot projects. We should encourage industries and localities to actively explore the construction of new technical facilities such as blockchain networks and privacy-preserving computing platforms. We should support the construction of data circulation and utilization facilities such as DSSNs and data components according to local conditions. We should support the construction of public service platforms for data circulation and trading. We should support the exploration and construction of infrastructure for cross-border data flows. We should establish data circulation access standards and rules and encourage the exploration of data circulation security technology, standards, and solutions.

(4) Build a convenient data delivery system

We should strengthen the design of the data exchange system and coordinate the optimized layout of data exchanges. We should support the innovative development of data exchanges and encourage all types of data to enter exchanges for trading. We should build an intensive and efficient data delivery infrastructure and provide a low-cost, highly efficient, and trustworthy data delivery environment for centralized on-exchange trading and decentralized off-exchange trading. We should promote the interconnection and interoperability of various data exchanges and trading platforms. We should promote technological innovations such as data value contribution assessment, data set recommendation and matching, and data product

difference analysis (差异性分析) and achieve accurate matching of supply and demand and convenient delivery. We should encourage local governments to improve their social service capabilities in data processing, testing, modeling verification, and security experiments and to create a "one-kilometer" working circle (“一公里”工作圈) that brings together industry, academia, research institutes, and users (产学研用).

(5) Build an industry data application system

We should strengthen the guiding role of scenarios, build a data application system for key industry fields such as industrial manufacturing, modern agriculture, digital finance, intelligent healthcare, smart transportation, cross-border logistics, shipping and trade, health, green and low-carbon industries, meteorological services, and digital culture, give full play to the main role of enterprises, and promote innovation in industry data applications. We should cultivate new products and services based on data factors of production, promote multi-scenario application of data and cross-entity reuse, and realize knowledge diffusion and value multiplication.

6. Compute base

(1) Promote the scientific layout of compute resources

We should accelerate the green development and organic coordination of diversified heterogeneous compute, such as general compute, intelligent compute, and supercompute. We should promote the clustering of various types of new compute in national hubs and nodes and strengthen the positioning of hubs and nodes into a national bastion of compute. We should build a national integrated compute network monitoring and dispatching platform. We should explore the use of storage-compute separation architectures to build new-style intelligent compute centers and new materials big data centers (新材料大数据中心).

(2) Promote the coordination of compute in the eastern, central, and western regions

We should strengthen the innovative application of emerging network technologies, optimize network billing methods, reduce data transmission costs between eastern and western regions, and promote the transfer of medium- and high-latency services from the east to the west. We should promote the interconnection and interoperability of compute, build a multi-level scheduling strategy engine for compute, achieve hybrid deployment and unified scheduling of compute resources across platforms, levels, and regions, promote the efficient connection of compute resources, and improve the efficiency of data aggregation, processing, circulation, and trading. We should promote 400G/800G high-bandwidth all-optical connections

between national hubs and nodes and demand areas, guide telecom operators and other entities to improve the efficiency of "public transmission channels," and promote the in-depth fusion of computing and networks.

(3) Promote the integration and innovation of compute, data, and algorithms

We should continuously expand the data compute (数算) industry ecosystem and help build a digital industry cluster with international competitiveness. We should promote the coordination of industry data and compute, realize trusted data circulation, and improve data processing capabilities and governance levels. We should establish and improve algorithm development and utilization mechanisms, actively carry out research on innovative algorithms and key technologies for large models, enhance data analysis capabilities, and reduce the level of compute consumption in large model computations.

(4) Promoting the integration of compute and green electricity

We should promote the transformation of the "watt" industry into the "bit" industry, strengthen the coordination and linkage between hubs and nodes and non-hubs and non-nodes (非枢纽节点), and support the integration of non-hubs and non-nodes with abundant green electricity resources into the construction of a national integrated compute network. We should strengthen the coordination and linkage between large-scale wind and solar bases and compute hubs and nodes to convert green electricity into green compute power. We should actively promote the consumption of wind, solar, and green electricity resources and help achieve carbon emission peak and carbon neutrality. We should support the use of new types of electrical power system models such as "source-grid-load-storage." We should strengthen smart energy management of data centers, carry out energy consumption monitoring and analysis and load forecasting for data centers, and optimize the overall operating efficiency of data center electrical power systems. We should explore new models of direct supply of green electricity and launch green electricity and green certificate trading in an orderly manner.

(5) Promote the coordination of compute development and security assurance

We should promote the construction of a national compute network foundational security service assurance platform and create integrated security assurance service capabilities. We should create a network and data security attack and defense exercise range and promote regular network and data security attack and defense exercises in national hub and node areas. We should build a testing site for compute network security application technology. We should strengthen the autonomous defense

capabilities of national hubs and nodes, achieve unified emergency response, unified security monitoring, and unified operation monitoring, and build security management measures for the entire life cycle.

7. Network support

We should build a high-speed data transmission network to achieve efficient and flexible data transmission and interconnection between different terminals, platforms, and private networks and solve problems such as insufficient data transmission capacity, high costs, and difficulties in interconnection. We should support telecom operators, relevant scientific research institutions, national big science installations (大科学装置), and other institutions in efforts to combine cloud-network integration technologies such as virtualized networking, network protocol innovation, and intelligentized task scheduling, form multi-party rapid networking and data exchange capabilities, and support elastic bandwidth (弹性带宽) and multi-dimensional billing (多量纲计费) for data transmission tasks.

We should promote the optimization and upgrading of traditional network facilities, advance the upgrading and evolution of 5G networks to 5G-A in an orderly manner, and comprehensively promote the R&D and innovation of 6G network technology. We should evenly distribute international telecommunications service entry-exit bureaus[2] in the eastern, central, and western regions and accelerate the expansion of international submarine and land cable information channels. We should build deterministic networks with guaranteed latency, stable bandwidth, and reliable transmission quality. We should deploy an "integrated space–ground" satellite Internet.

8. Security protection

The focus of the construction of the national data infrastructure security assurance system is to build a multi-level, omnidirectional, and three-dimensional national data infrastructure security assurance framework that runs through the entire data life cycle, helps all participants improve their data security assurance capabilities, and ensures the credibility, integrity, and security of the data.

At the national data infrastructure security level, we should achieve security management such as trusted access, secure interconnection, cross-domain

[2] Translator's note: International telecommunications service entry-exit bureaus (国际通信业务出入口局), also known as international telecom entry-exit points (国际通信出入口), connect the Chinese internet to the global internet. China established a total of nine of these in the cities of Beijing, Shanghai, and Guangzhou in 1994, and approved the construction of additional ones in the cities of Nanning, Qingdao, Kunming, and Haikou in 2024.

management and control, and full-stack protection, establish dynamic discovery, real-time alerts, comprehensive analysis, collaborative processing, and cross-domain protection from network security risks and threats. Traceability and situation control capabilities provide endogenous protection capabilities against security threats such as built-in backdoors and vulnerabilities in chips, software, hardware, and protocols. We should strengthen prevention of and response to the internal data security risks of cooperation partners, operations and maintenance (O&M) personnel, and platform users. We should strengthen emergency response to external threats such as intrusion and penetration, denial of service, data theft, ransomware, and data poisoning.

In terms of data circulation and utilization security, we should comprehensively utilize technical means such as privacy-preserving computing, blockchain, and data usage controls to ensure the trusted collection, encrypted transmission, reliable storage, controlled exchange and sharing, destruction confirmation, and evidence preservation and traceability of data, etc. Avoid risks such as data privacy breaches, violation of regulations, and abuse. We must strengthen security audits of algorithms, models, and data, enhance model robustness and security, ensure that high-value, highly sensitive data are "available and invisible," "controllable and measurable," and "traceable and auditable," ensuring security throughout all stages of the data life cycle.

9. Organizational assurance

(1) Improve the policy assurance system

We must establish and improve the data infrastructure system, and accelerate the promulgation of policy documents on data property rights, circulation and transactions, income distribution, and security governance. Under the new infrastructure (新型基础设施) planning arrangement, we should research and formulate a national data infrastructure construction plan. We should increase central government investment support for national data infrastructure construction. All regions and departments should provide key support in data infrastructure planning and layout, funding arrangements, and research projects. We should actively channel the power of social capital[3] to participate in the national data infrastructure construction.

(2) Accelerate technological innovation and exploration

We should support qualified industries and regions in carrying out pilot trials and explore the construction of data infrastructure. We should encourage enterprises and

[3] Translator's note: The Chinese term 社会资本, translated literally as "social capital," refers to any source of funding outside of government budget outlays. This term encompasses investment by private individuals and private institutions. However, investment from state-funded entities such as state-owned enterprises (SOEs), including state-run banks, also falls under the umbrella of "social capital."

scientific research institutions to increase R&D investment, speed up the research and development of key technologies for data circulation and utilization, and accelerate the conversion of major achievements into practical applications. We should promote technological innovation through the establishment of National Key Research and Development [Program] projects and topics, winner-takes-all open competition,[4] and data technology innovation competitions.

(3) Strengthen standards and talent support

We should strengthen standards support and research and formulate relevant standards and specifications for data infrastructure. We should encourage enterprises, community groups (社会团体), and scientific research institutions to participate in the formulation of international standards for data infrastructure. We should strengthen cooperation with international standardization organizations such as the International Organization for Standardization (ISO), International Electrotechnical Commission (IEC), International Telecommunication Union (ITU), Institute of Electrical and Electronics Engineers (IEEE), and 3rd Generation Partnership Project (3GPP) and encourage high-level experts in the data field to serve in international organizations. We should promote talent team building and establish data talent evaluation standards and selection mechanisms.

Appendix: Explanation of technical terms

(1) Data circulation and utilization technology

In the field of data circulation and utilization, the commonly used technical approaches currently include privacy-preserving computing, blockchain, and data usage controls.

1. Privacy-preserving computing

Privacy-preserving computing refers to a type of information technology that analyzes and computes data under the premise of ensuring that the data provider does not leak the original data. This ensures that the data is "available but invisible" in all stages throughout the data circulation process, including generation, storage, computation, application, and destruction. Common technical solutions for privacy-preserving computing include secure multi-party computation, federated learning, trusted execution environments, and confidential computing. Common underlying

[4] Translator's note: The idea behind "winner-takes-all open competition" (揭榜挂帅), in the context of Chinese science and technology projects, is that the government openly lists the technological breakthrough(s) it desires. Any individual or group in society, not just a select few, are then eligible to win a cash award if they succeed in making the breakthrough.

technologies include garbled circuit, oblivious transfer, secret sharing, and homomorphic encryption.

2. Blockchain

Blockchain is a new type of database software that integrates multiple technologies such as distributed networks, encryption technology, and smart contracts. It has the characteristics of decentralization, trusted consensus, immutability, and traceability. Blockchain is mainly used to solve trust and security issues in the process of data circulation.

3. Data usage controls

Data usage controls refer to the use of technical means to control the transmission, storage, use, and destruction of data. For example, through smart contract technology, the data usage intentions of data rights and interest holders are converted into machine-readable smart contract terms, solving the precondition problem (前置性问题) of data controllability and achieving control over factors such as the times, places, subjects, behaviors, and objects of data asset use.

(2) Practical solutions for data circulation and utilization

In the data circulation and utilization field, the industry's current practical solutions mainly include trusted data spaces, data exchanges, DSSNs, and data components.

1. Trusted data space

A trusted data space is a data circulation and utilization facility based on consensus rules. It connects multiple parties to achieve the sharing and use of data resources. It is an application ecosystem for the co-creation of the value of data factors of production and an important vehicle that supports the construction of a national integrated data market. A trusted data space must possess three core capabilities: trusted data management and control, resource interaction, and value co-creation.

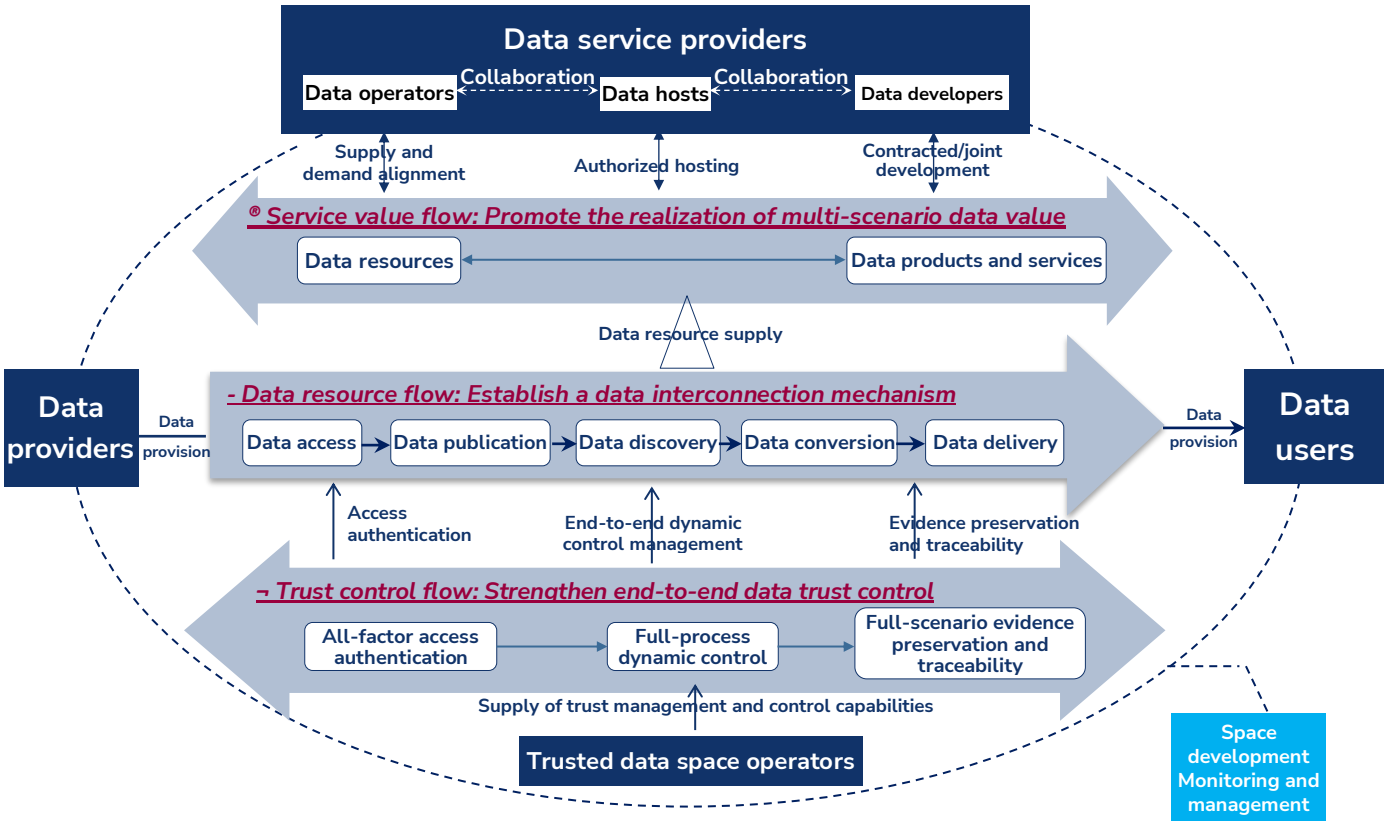

Figure 2. Trusted data space architectural diagram

2. Data exchanges (数场)

Data exchanges are comprehensive data circulation and utilization facilities based on open networks and compute, and on privacy-preserving computing, blockchain, and other related functional facilities. Data exchanges provide online and offline resource registration, supply-demand matching, trading and circulation, development and utilization, and evidence preservation and traceability for data factors of production. They are a type of comprehensive data circulation and utilization facility that supports multi-scenario applications. A data exchange aims to provide efficient circulation, value unleashing, and a vibrant ecosystem. It has the ability to make data visible, accessible, usable, controllable, and traceable, and features openness, integration, and scalability.

A data exchange constructs a standardized technical framework based on five dimensions: points, lines, surfaces, fields, and security. A point is an access point by which a data subject (数据主题) enters a data exchange. Lines represent high-speed data transmission networks that connect various entities and platforms within the data exchange, achieving interconnection among various entities in the digital exchange. A surface is a collection of data entities and transmission networks in a data exchange. It is the core unit for achieving large-scale data circulation and efficient and secure data utilization. Data exchange infrastructure is constructed through the combination of

points into lines and lines into surfaces. An exchange (场) is a general term for data applications, scenario-based innovations, and related capabilities, processes, and specifications built on data exchange infrastructure. Security is the dynamic and end-to-end protection measures covering points, lines, surfaces, and exchanges. The technical architecture of a data exchange includes access points, functional platforms, management platforms, security assurance, network transmission, and other foundational service platforms.

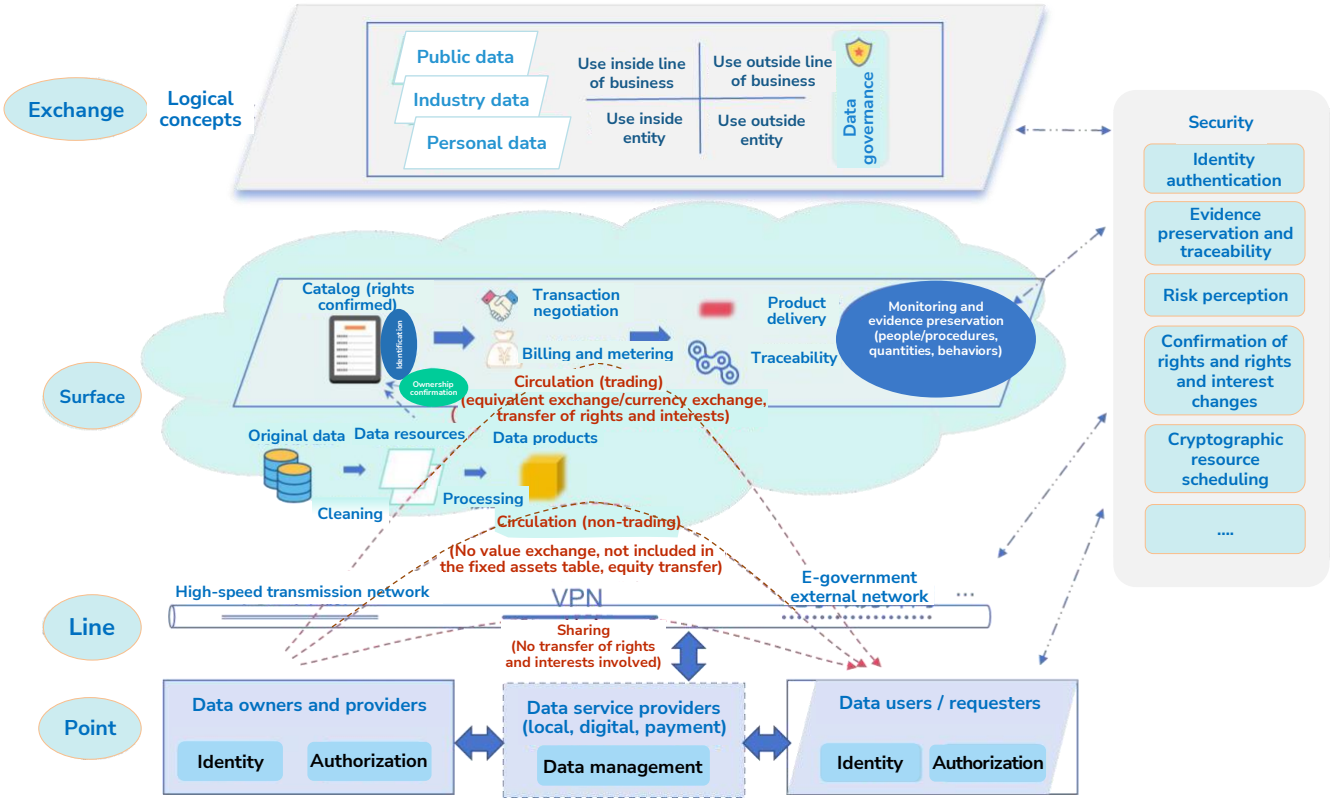

Figure 3: Functional architecture of a data exchange

3. Data switching service network (DSSN; 数联网)

A DSSN consists of data circulation access terminals, data circulation networks, and data circulation service platforms, providing data circulation services with one-point access, wide connectivity, standard delivery, security and trustworthiness, compliance supervision, and openness and compatibility.

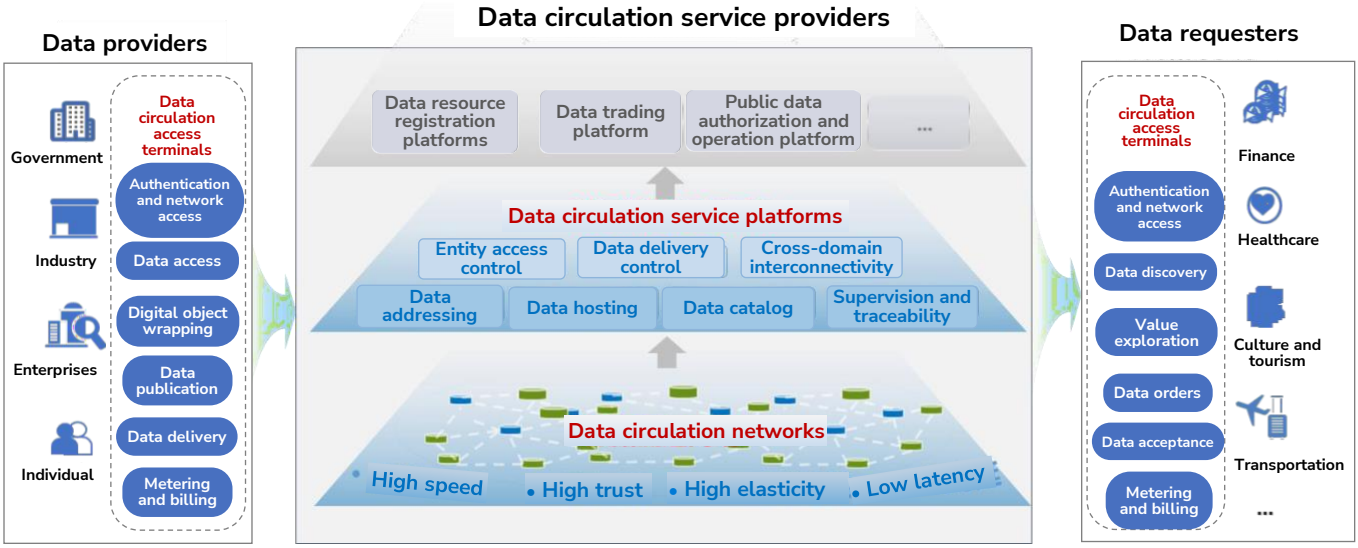

Figure 4: Functional architecture of a DSSN

4. Data components

Data components provide data storage and processing services that feature unified standards, independent controllability (自主可控), security, reliability, and end-to-end supervision. They support the use of standardized processes to complete large-scale processing, production, and reuse of data products and are suitable for large-scale data processing and production scenarios. As the "intermediate state" connecting the two ends of data supply and demand, data components "decouple" (“解耦”) the original data from data applications. Based on data component-related modules, the development and management of data factors of production is implemented throughout the entire life cycle from data collection to data component processing and transactions.

(3) Data security technology

Data security technology provides security protection for the entire life cycle of data, from data collection to storage, processing, transmission, sharing, and destruction, including data backup and recovery, application data encryption, data leak detection, flow monitoring, identity authentication and access control, data masking, data watermarking, and data security situational awareness.