



中国科学院计算机网络信息中心
Computer Network Information Center,
Chinese Academy of Sciences



e-Infrastructure for Future

China SRC

CNIC,CAS

Prof. Jianhui Li

22rd Mar. 2024, Shanghai, China



Outline

- CNIC introduction
- Understanding of SRCNet
- Views on China SRCNet
- Our current and future capabilities

Computer Network Information Center, CAS

Birthplace of China's Internet: China's first router, first domain name service, access point of full-functional connection between NCFC and NSFnet

Pioneer of HPC: China's first public supercomputing service for science

Contributor to open science and management informatization



Campus Distribution



CAS Informatization Plaza



Building at Software Park of CAS



Park at Huairou Science City

Hello Science | Hello World

Birthplace of China's Internet

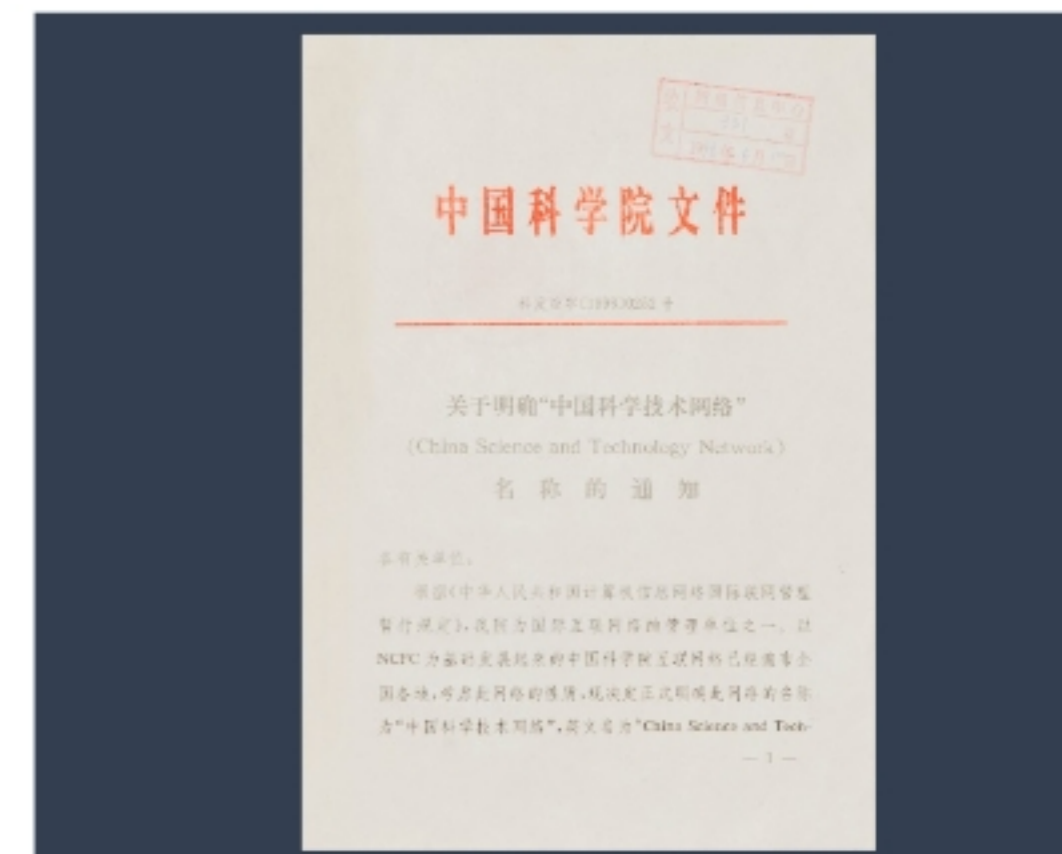
China became the 77th country on the Internet with a 64 Kpbs full-functional connection between National Computing and Networking Facility and NSFnet in 1994. The access point was located at the founding place of CNIC, CAS. Established in 1995, CNIC is a research and service center to support CAS digitalization related to research, management and public outreach, demonstration base for IT innovation and applications.



first router in China, 1994



first CN domain name server
in China, 1994



National Research Network
CSTNet, 1996-

High Performance Computing

CNIC provided China's first public supercomputing service to support science in 1996, and develops high performance computing environment and algorithm for efficient implementation on supercomputers.



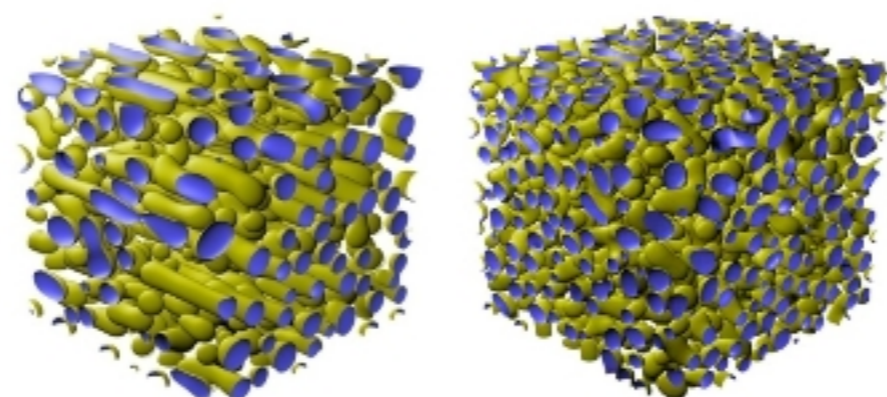
1996, SGI Power Challenge XL



2014-, Yuan Supercomputer



2019-, ORISE



ACM Gordon Bell finalist

Extreme-Scale Phase Field Simulations of Coarsening Dynamics on the Sunway TaihuLight Supercomputer

Scientific Data and Public Outreach

CNIC supported data initiatives of Chinese Academy of Sciences since its foundation, and develops **general-purpose technologies, standards and services** to promote the availability and usability of data for all areas of research. Based on early databases, CNIC **opened its resources to the general public** and initiated Virtual Science Museums of China in 1999.



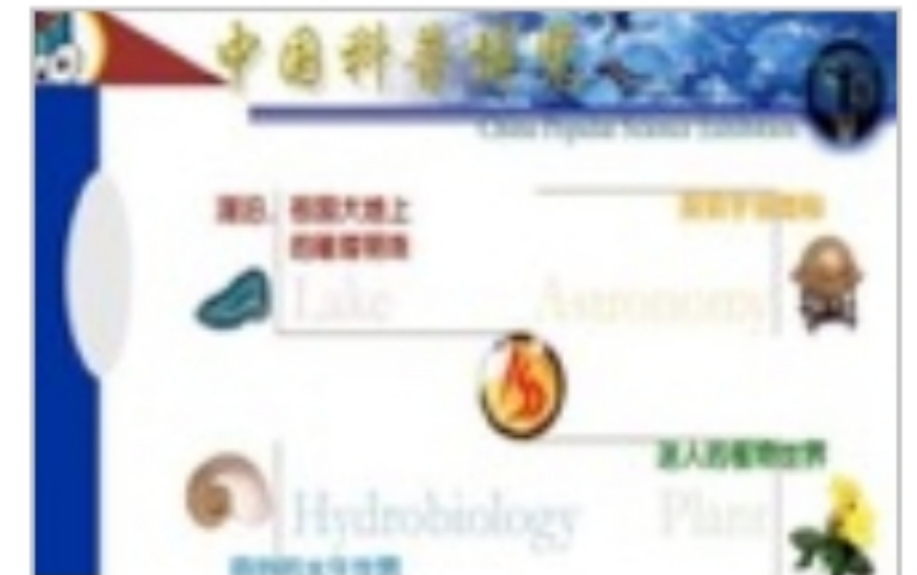
CODATA China Secretariat
(1984) 1995-



General Data Center, CAS
(1986), 2019-



National Basic Science Data Center
(1986), 2019-



science outreach, 1999-

International Cooperation

- o Host organization of CODATA China secretariat, Co-Chair organization of APAN China
- o Founding member of GLORIAD, PRAGMA, GOSC, EARBN
- o Cooperation with ISOC, ICANN, GEANT, EGI, DataCite, KISTI, NII, NICT, Figshare, Springer Nature, AICAD, UCSTI, etc.
- o Experts holding EC or AC positions in CODATA, APAN China, W3C China



Outline

- CNIC introduction
- Understanding of SRCNet
- Views on China SRCNet
- Our current and future capabilities

What is SRCNet?

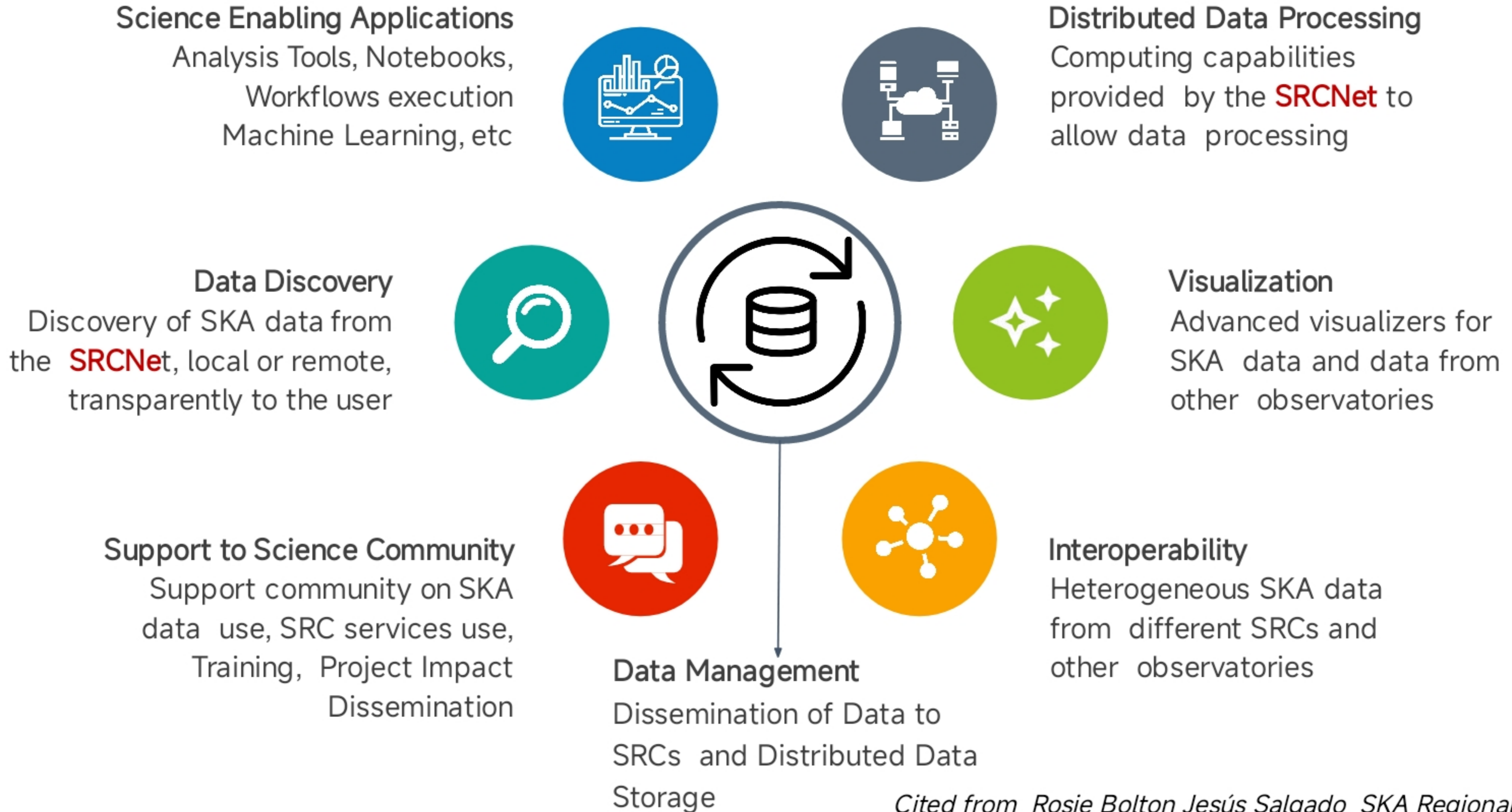
SRCNet is an international network of interoperating science centres.

SRCNet mission: To ensure that scientists can access SKA data products and use them to make discoveries.

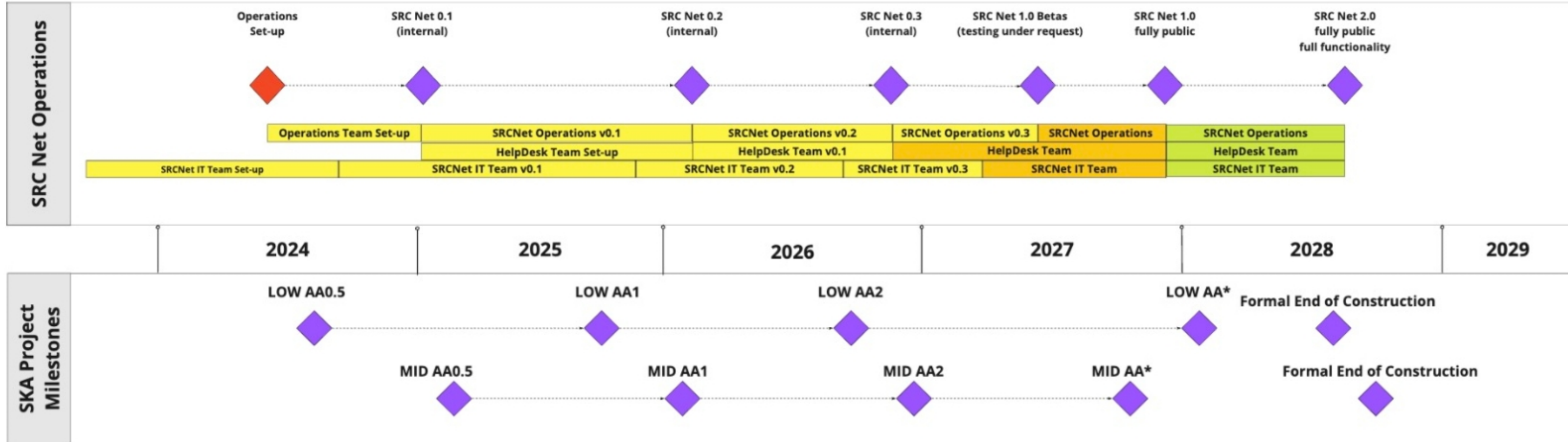


- Collectively meet the needs of the global community of SKA users
- Anticipate heterogeneous SRCs, with different strengths

SKA Regional Centre Capabilities



SRCNet 0.1 Current Status



Set-up during 2024; Get started at 2025; Finished by early 2026;

SKA Regional Centre Challenges



100 Gbps



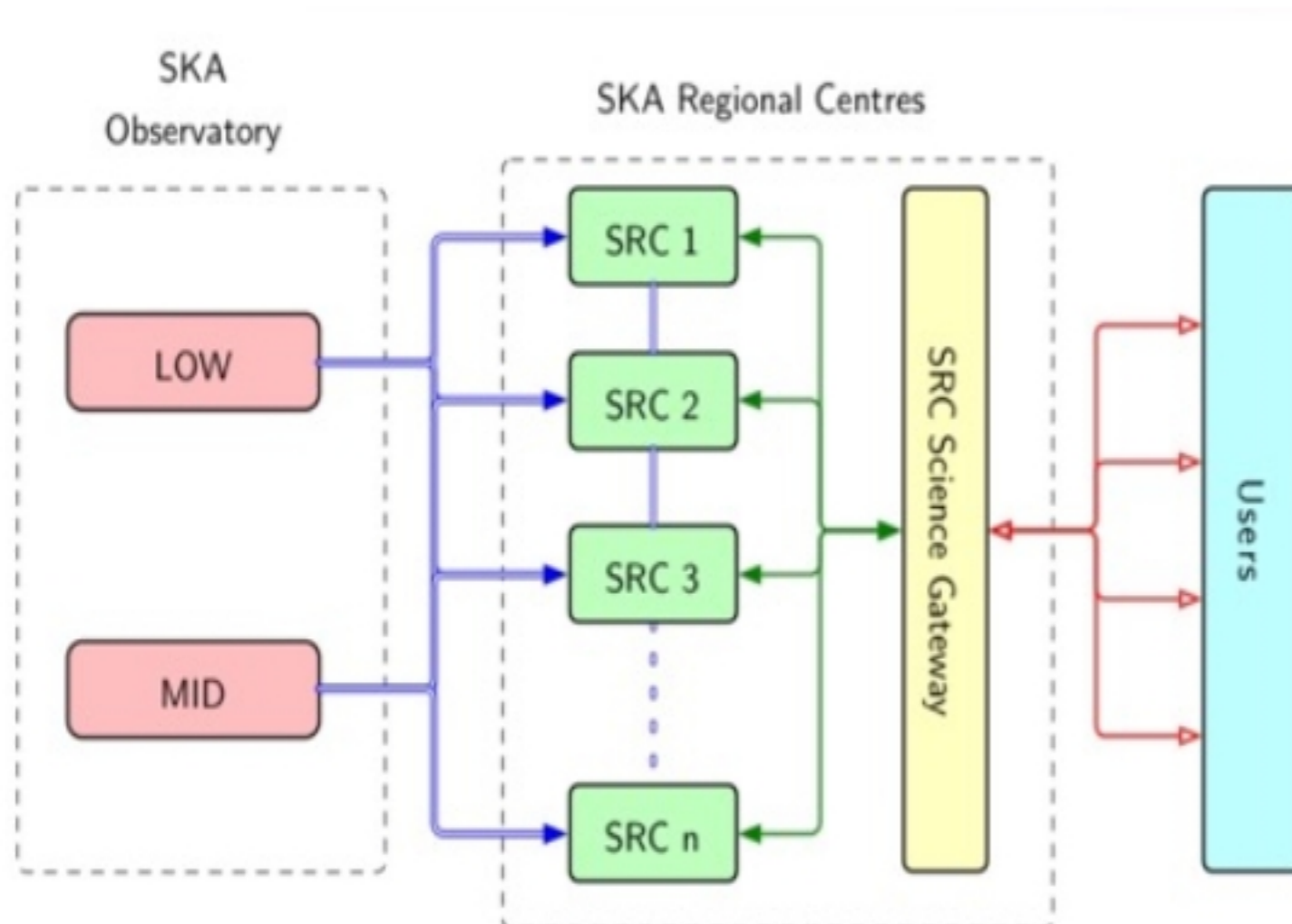
300 Pflops



100 PB/year



software tools



Outline

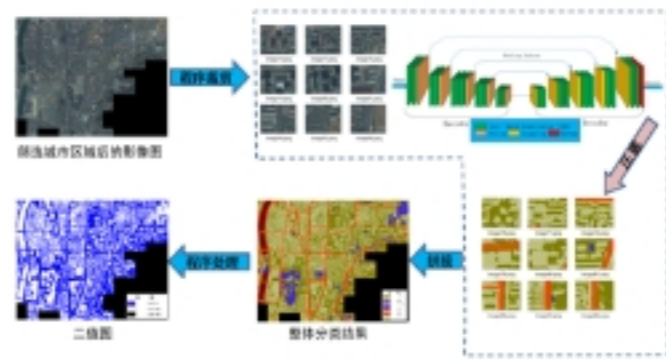
- CNIC introduction
- Understanding of SRCNet
- **Views on China SRCNet**
- Our current and future capabilities

Open Science is a New Paradigm of Global Research Innovation

To make full use of the world's scientific and technological resources to address the major challenges facing humanity, UNESCO has launched and promoted the Open Science Campaign



Carbon neutrality, carbon peaking program process



Heterogeneous, complex environment data analysis



Embracing global restructuring in the post-COVID-19 era



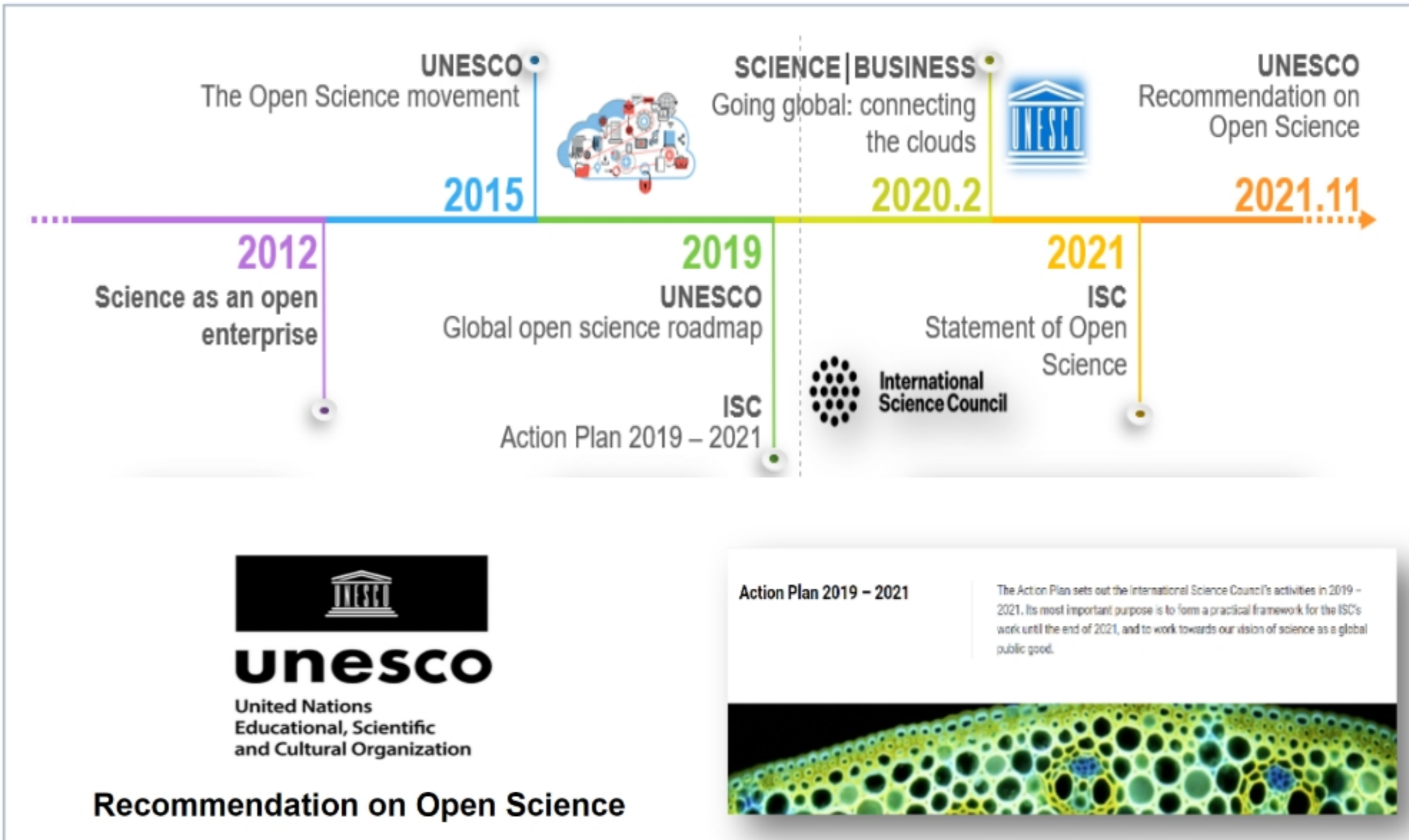
Long-term work on population and health



Urgent need of technological innovation iteration

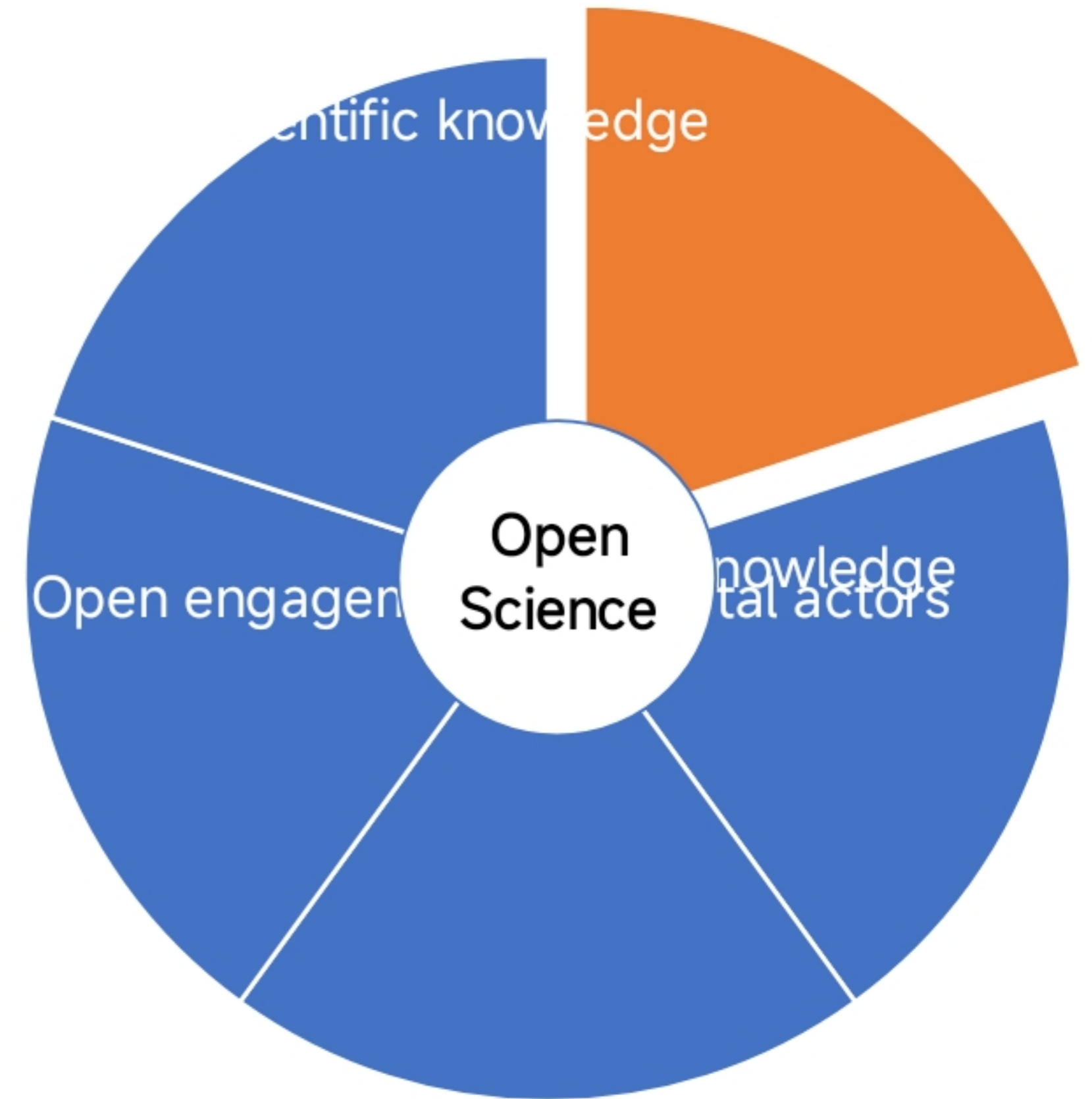


United Nations Sustainable Development Goals



Open Science and Open Infrastructure

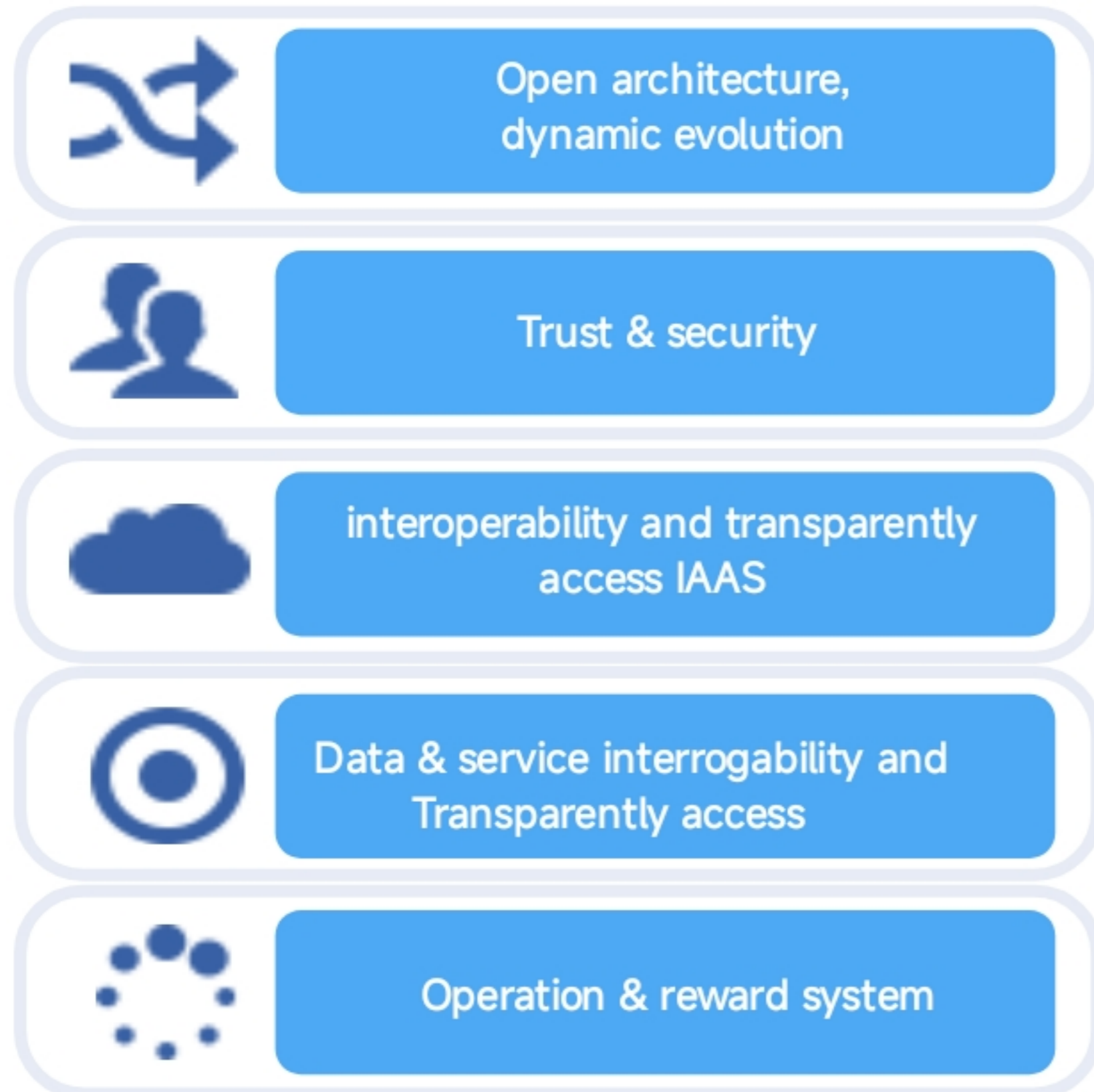
- Open science infrastructure is rich in connotation and diverse in types
 - ✓ Research network, computing, storage; Open research data and software; Open access platform; virtual lab...
- Open science infrastructure has the characteristics of federation, accessibility, interconnection and interoperability
 - ✓ Open science infrastructures should be federated, accessible, internationally interconnected, interoperable (UNESCO, 2021).
- The global cooperation of open science infrastructure is the inevitable way to realize the flow and sharing of global open science resources and services



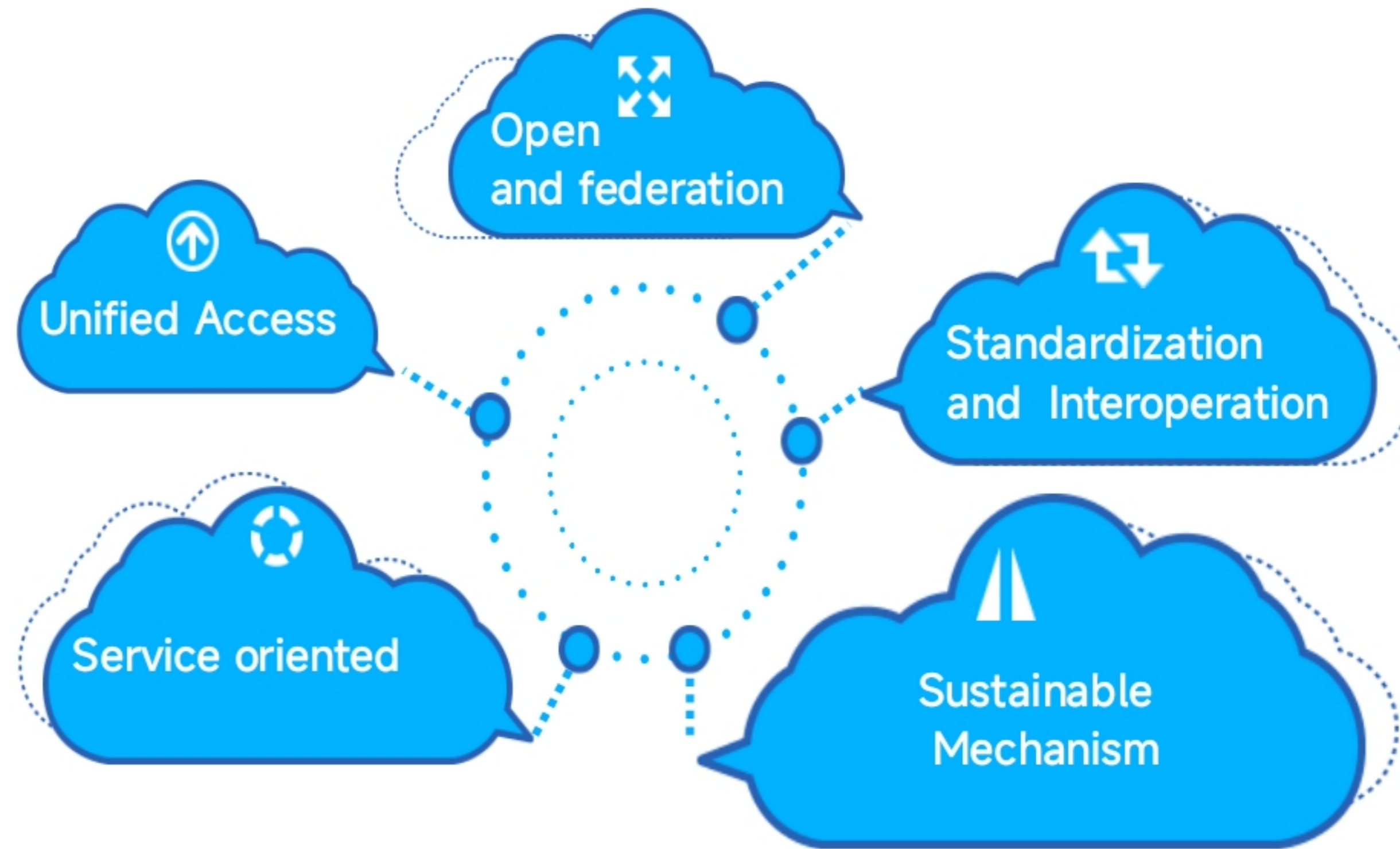
**FIVE KEY PILLARS OF OPEN SCIENCE
(UNESCO OPEN SCIENCE
RECOMMENDATION, 2021)**

Integrate Distributed e-Infrastructure for Open Science

- Open science infrastructure features: **Federated, Accessible, Internationally Interconnected, Interoperable** (UNESCO, 2021).



challenges



principles

Our Perspective on SRCNet in China

- Building a centralized data center to meet SRC indispensable needs
- Making full use of existing e-infrastructure capabilities
 - Network: CSTNet and Partner
 - Computing: CNGrid, CASGrid, National Supercomputing Center
 - Storage: CSTCloud Storage backup and archive
 - Software: AAI, data management, data transmission
- Constructing an Open and federated e-infrastructure with physical distribution and logically unified data layer

Outline

- CNIC introduction
- Understanding of SRCNet
- Views on China SRCNet
- **Our current and future capabilities**

Requirements for e-Infrastructure in CAS

Data-intensive & big data driven Science

Major Research Infrastructures



National scientific data centers



Field Stations all around China



Major international cooperation projects



- Decades investment in cyber-infrastructure development and e-science in CAS, totally over 2 billion RMB

CSTCloud: A national research e-infrastructure

- CSTCloud supports multidisciplinary open scientific researches with integrated cloud services for the discovery, usage and delivery of S&T resources.

- **Continually funded by CAS Informatization Program(2017-2020,2020-2025)**

2016.12 The State Council issued the National Informatization Plan for the 13th Five-Year Plan, "CSTCloud" was included in the priority action



2018.04 CSTCloud portal launched

2019.04 CSTCloud Phase II Construction Plan Passed the review (focus on convergence capability and service)



2019.12 CSTCloud 2.0 launched



2016.12

2017.01

2017.12

2018.04

2018.11

2019.04

2019.07

2019.12

2020.6

2020.11

2017.01 The Chinese Academy of Sciences issued the "13th Five-Year Plan for Informatization Development of the Chinese Academy of Sciences, establishing the task of "CSTCloud"

2017.12 Launching Ceremony of "CSTCloud"

2018.11 CSTCloud Phase I Construction Task, passed the review (focus on resource performance construction)

2019.07 Resource Aggregation Workshop

2019.06 MoU Signed between CNIC & EGI.eu, mobilizing and co-design of GSOC Initiative

2020.11 GOSC workshop held only jointly by CNIC, EGI and CODATA.

CSTCloud Service Catalogue and System

AI Operation and Maintenance System

Security Monitoring and Protection System

National Large Research Infrastructure
National Science Data Center
National Field Scientific Observation and Research Station

Space, Astronomy, High Energy Physics,
Microbiology ... Thematic Cloud

Information Technology Innovation and
Research
Demonstration Application

SAAS



New ARP



Website
Group



Science
Communication



CAS
Mail System



Cloud
Conferencing



Cloud
Box



Conference
Service Platform



Overleaf



Open Source
Community

PAAS



CSTCloud
Passport



CSTCloud
AAI



Jupyter
Hub



Cloud
Database



CAS Science
Data Center



National & CAS Grid



National Basic
Data Center



eduroam CN
Root

IAAS



Domestic and International Network



中国科技网
China Science & Technology Network



HPC & AI Heterogeneous System

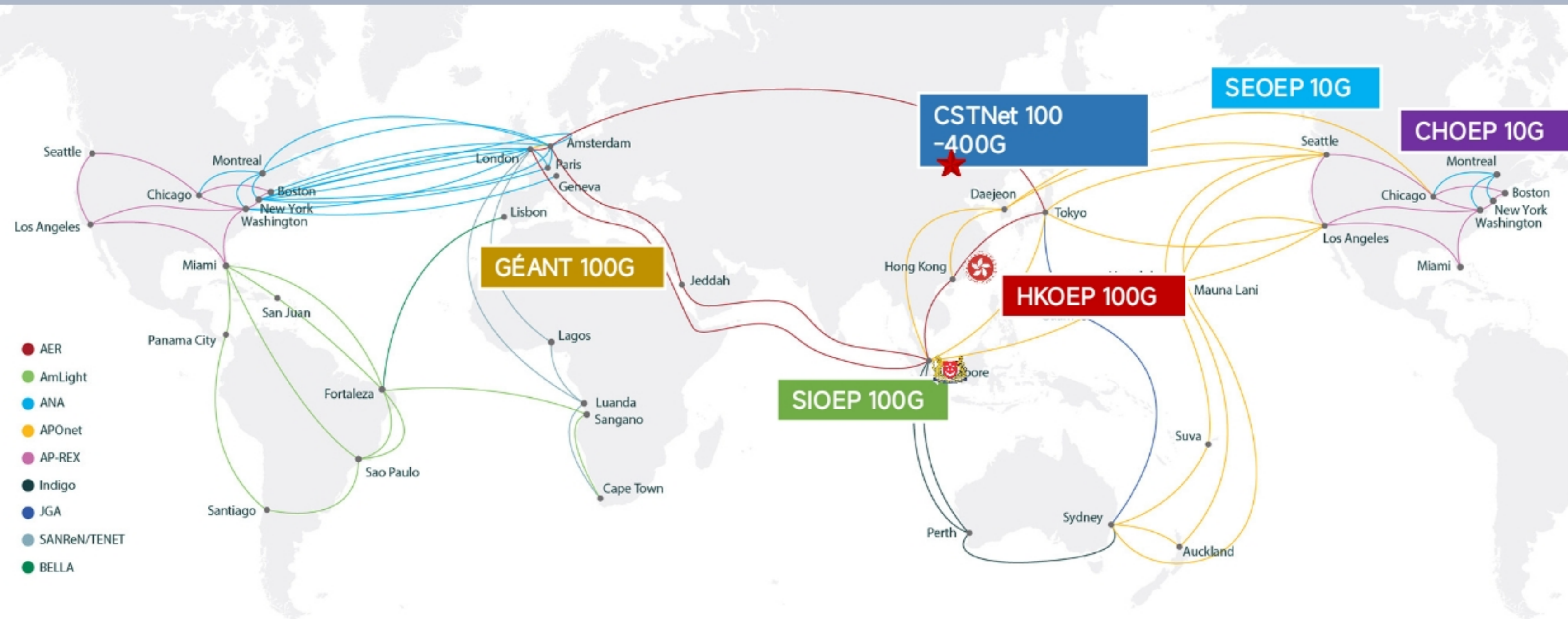


Federated Cloud and Storage



CSTNet Has Become Part of Global 100G NRENs

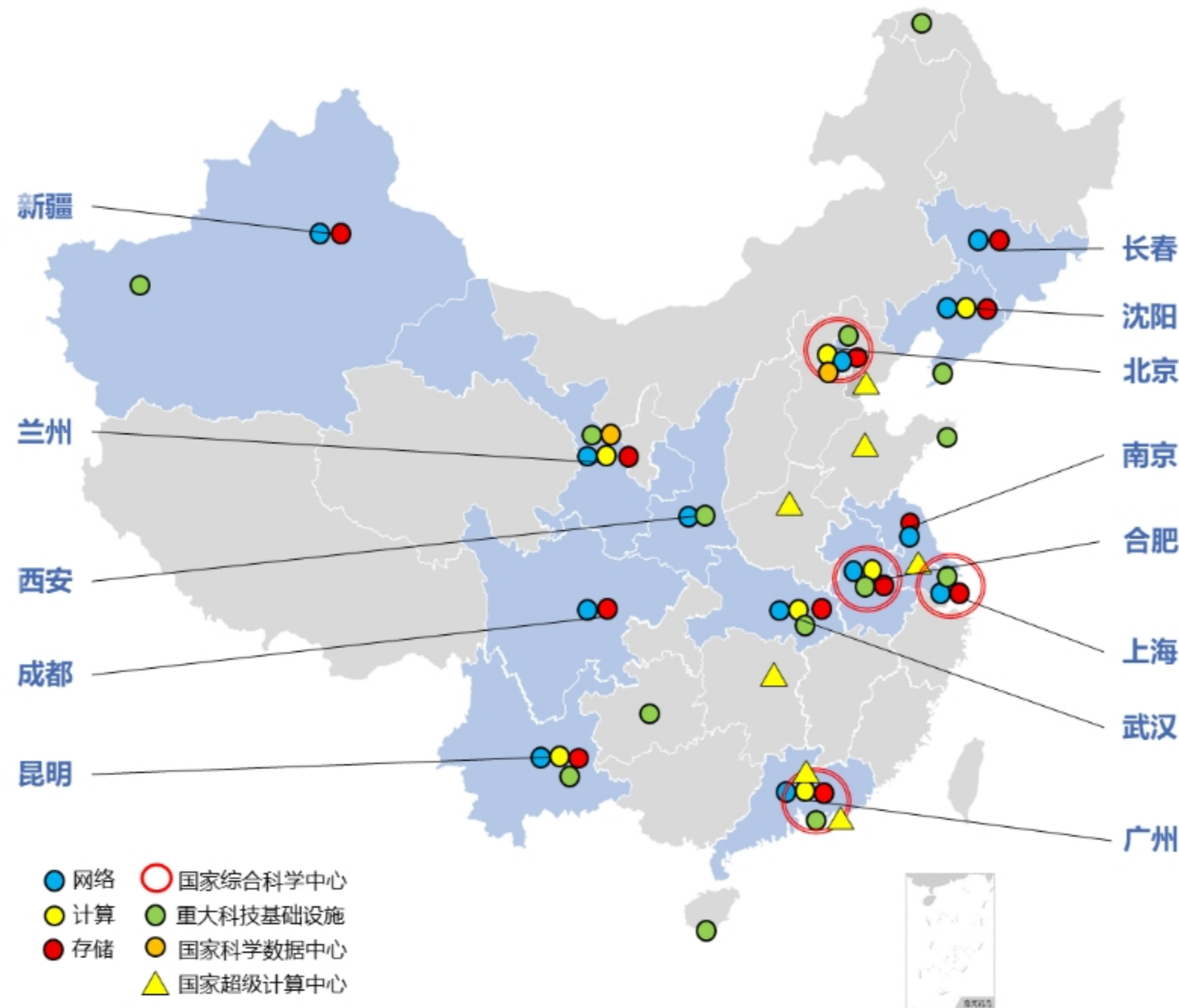
GLORIAD → HKOEP → OrientPlus → Seattle and Chicago Ex Point → China-EU 10G low-latency terrestrial link → SIGOEP → EU-China 100G



Connected Ecosystem for Digital Transformation

CAS research, management and collaboration supported by information infrastructure and services converging network, storage, computing, and information resources

- 4 national science centers
- 21 national infrastructures
FAST, LHAASO, EAST.....
RIs in Huairou, Zhangjiang, Hefei
- 4 international cooperation
SDGs, LHC, SKA, ITER
- 44 national observatories
and field stations
- 10 CAS informatization projects



3 international research innovation centers

55+ key projects
CAS Earth Big Data Service Platform, eVLBI, ground transmission network service for Space Pilot Program, Space Science Satellite Data Service, etc.

19 science data centers

N CAS info systems

"Global Open Science Cloud" (GOSC) Initiative



in 2019, we proposed and launched the "Global Open Science Cloud" Initiative, aiming to achieve interconnection, interoperability and federal services of global open science platforms. Currently, members include more than 200 scientists from 41 countries or regions.



GOSC is highly recognized by the international community:

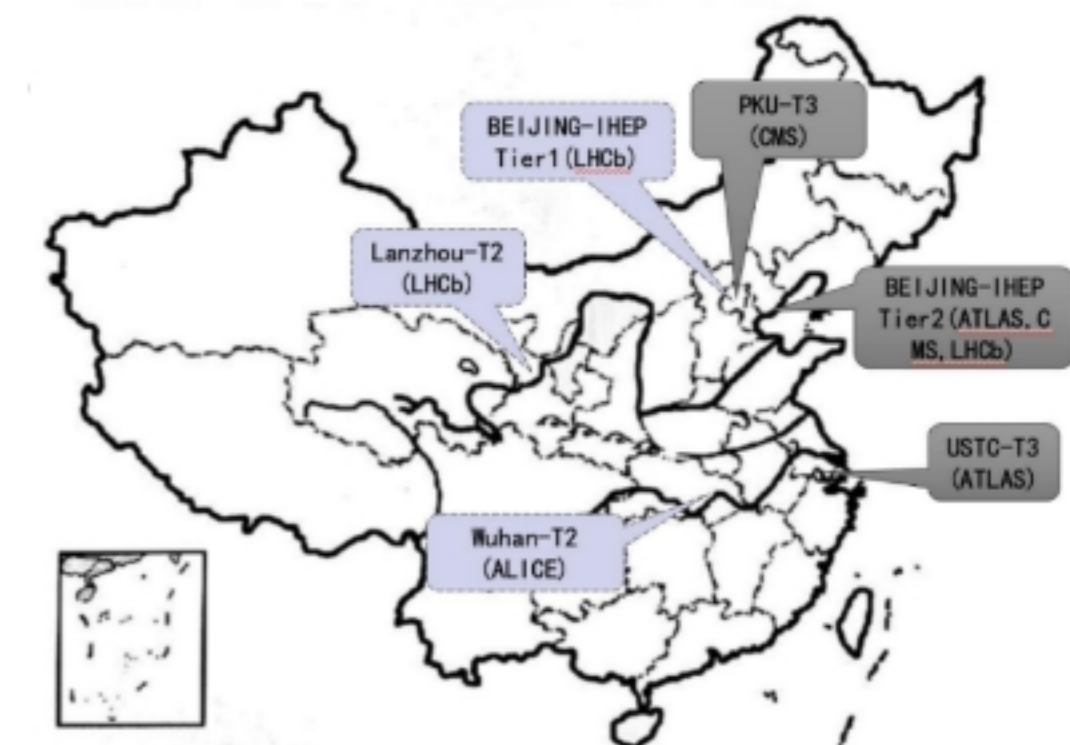
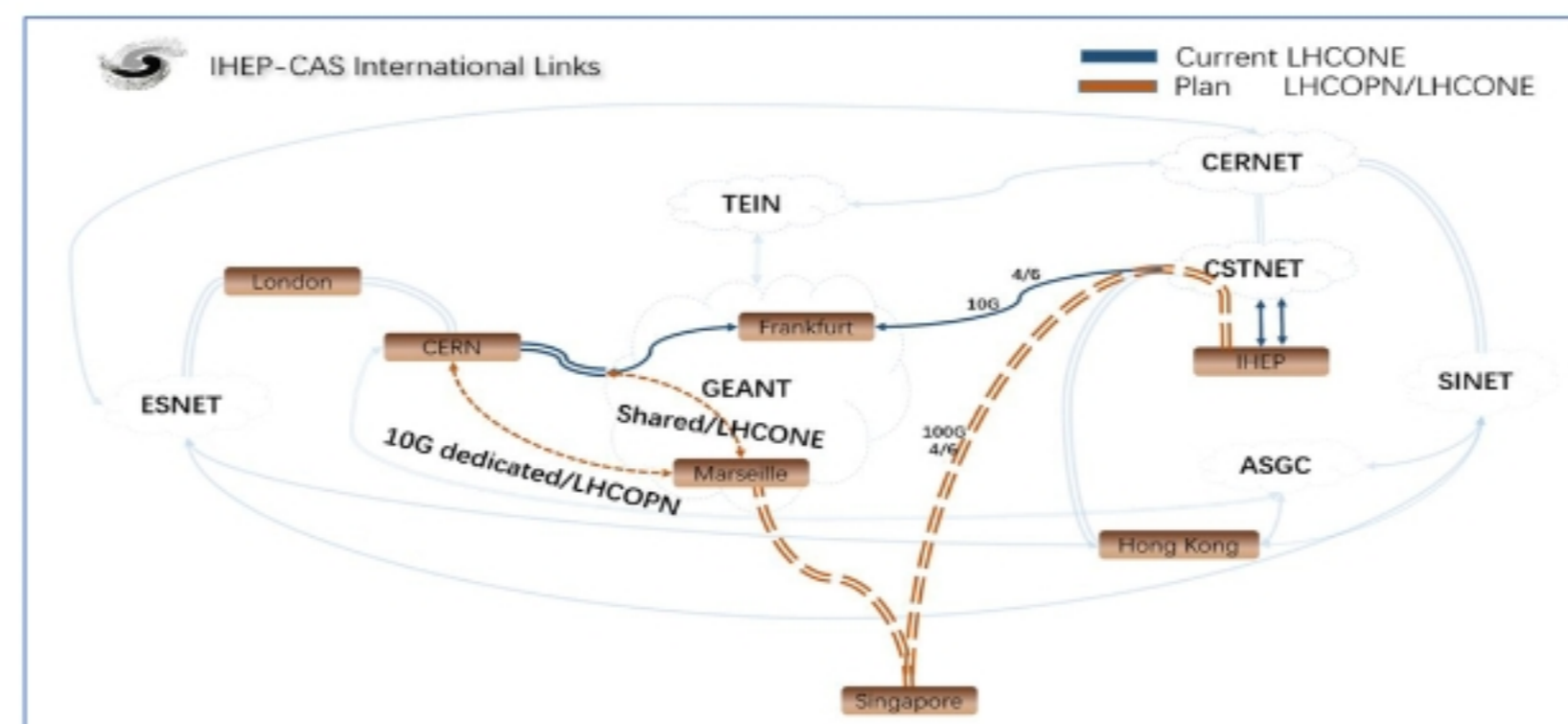
- The UN believes that GOSC has made outstanding contributions to the realization of SDGs goals in African and Arab countries.
- ISC names GOSC as key milestone in 2023 Open Science mission to serve the Global South
- CODATA lists GOSC as one of the three priority development goals in its ten-year development plan

The establishment of GOSC IPO is CODATA's first global international project office since 1966.



Typical Case 1 : CERN LHCONE and LHCb Tier1

- 2018.4, launched LHCONE
- 2019.9, launched 4X10G dedicated lines based MPLS VPN
- 2023.10, launched a 20Gbps virtual dedicated line to CERN, helping IHEP to upgrade to LHCb Tier1, supporting the processing and distribution of 6~8% of LHCb's original data, and a massive scientific research data of 10PB/year

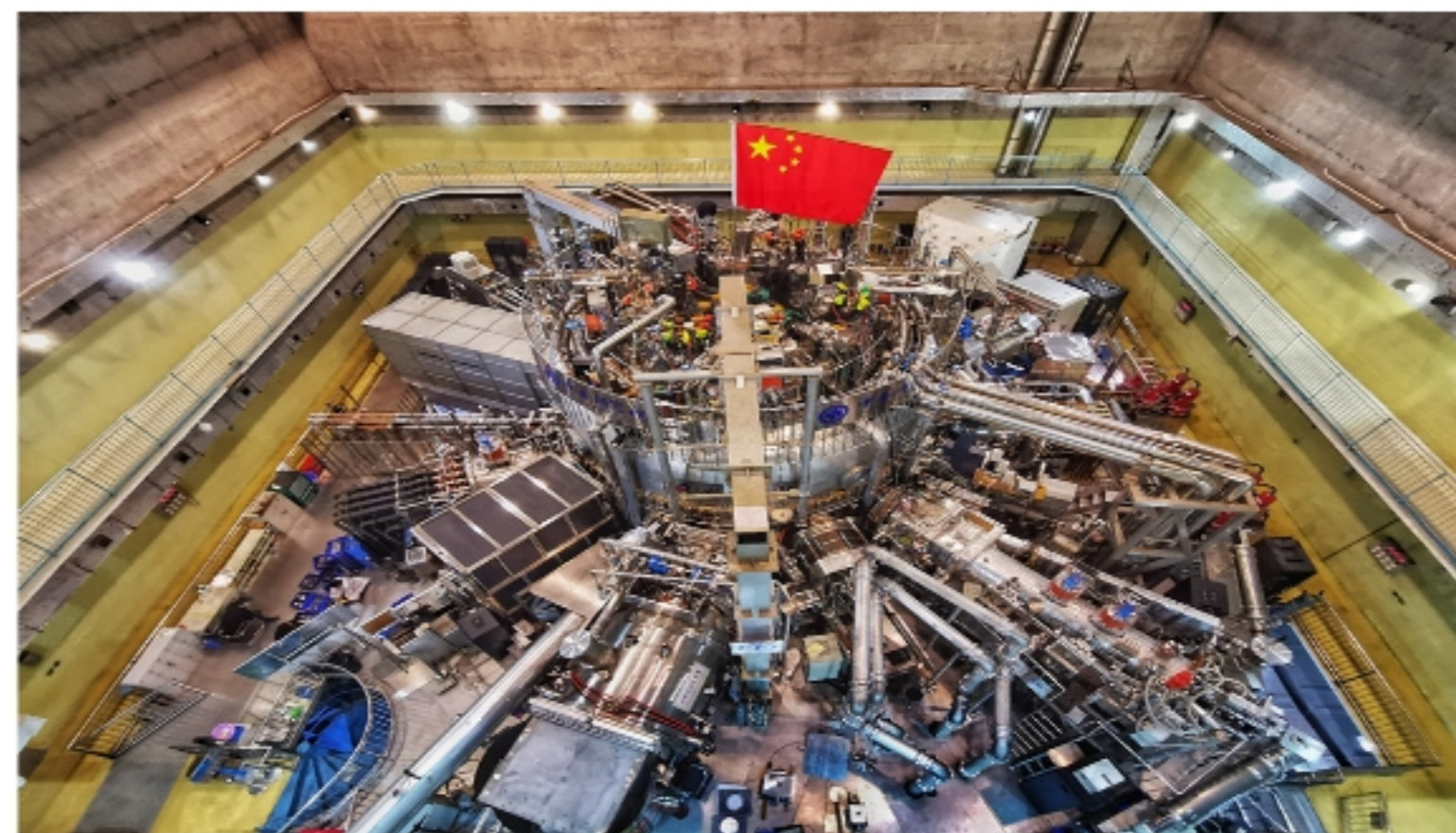


Typical Case 2 : ITER & EAST

- CSTNet conducted network research, design, testing and construction based on the characteristics of ITER's application requirements.
- Build a robust and scalable IT platform to lay a solid foundation for ITER's scientific output.

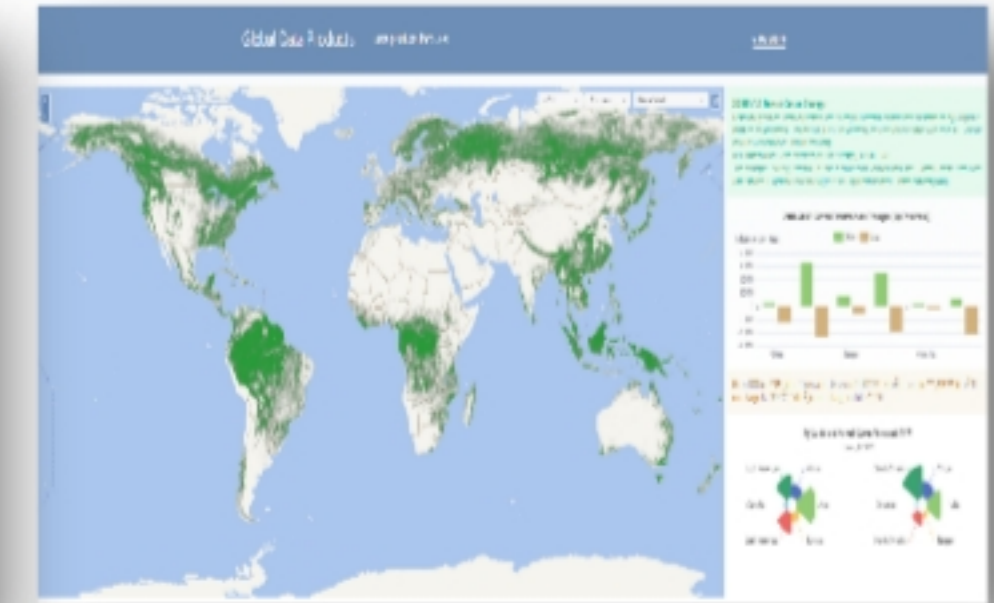
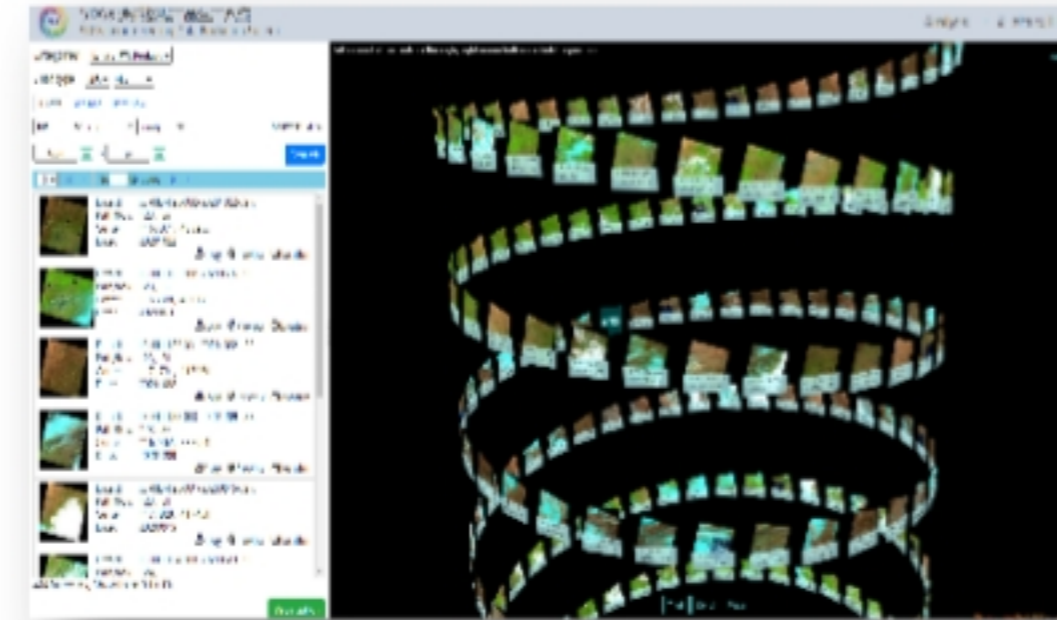


- CSTNet provides international and domestic high-speed network transmission service for EAST .
- CSTCloud provides storage backup service for EAST data center.



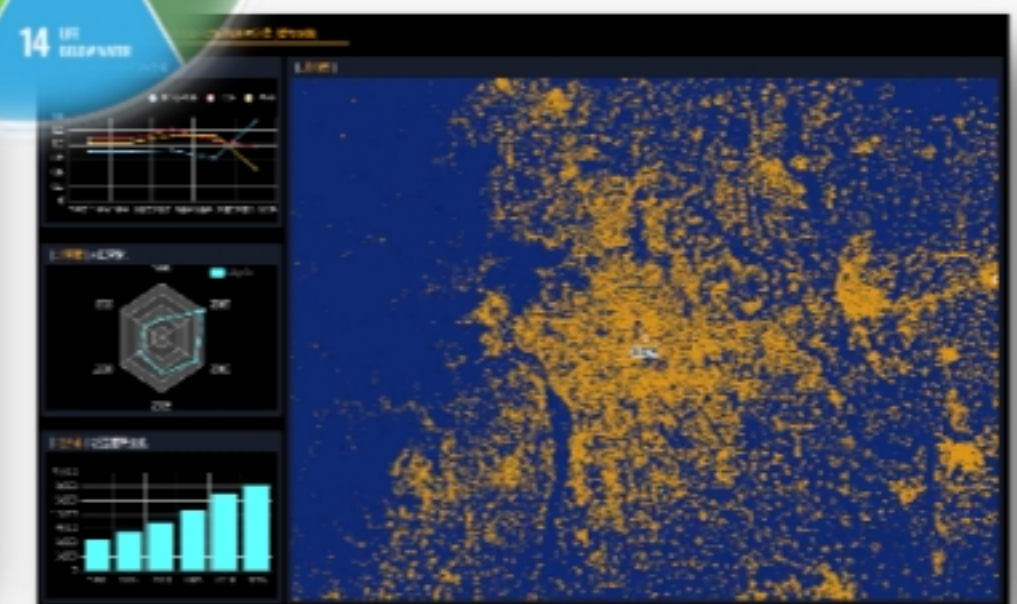
Typical Case 3 : CSTCloud for SDGs

- The *SDG Big Data Platform* provides rich data resources and cloud services for SDGs Decision Making and science discovery, that
 - With **10PB data** covering multiple subjects, such as geography, remote sensing, ground monitoring, and social statistics;
 - **Characterize and Profile Scientific Workflows** featuring lifecycle Big Earth Data management, including massive data storage, curation, computation, analysis and visualization;
 - Integrate **over one hundred algorithms and tools** for advanced data analysis and management;
 - Provide **one-stop cloud services** through the **bilingual portals** to support the UN Sustainable Development Goals.



Databank Remote Sensing
Data Engine

SDGs Data Products



SDGs workbench

SDGs Decision Making

SDGs indicator on-demand computing

1. Upload data to the EarthDataMiner system

The screenshot shows the EarthDataMiner interface. On the left, a 'Browser data' table lists various datasets. An arrow labeled 'Upload data' points from this table to a central workflow diagram. The diagram consists of several steps: '指定以城市/时间范围的遥感影像数据准备' (Specify remote sensing data preparation by city/time range), '遥感影像中目标识别分析' (Remote sensing target identification analysis), '指标在线提取与展示' (Online extraction and display of indicators), '遥感影像产品 Landsat, Sentinel, MODIS GF1等' (Remote sensing products: Landsat, Sentinel, MODIS, GF1, etc.), '数据的预处理' (Data preprocessing), '按照区域进行数据叠加' (Data overlay by region), and '按照公式进行指标计算' (Indicator calculation by formula). A 'Task S' icon is visible on the right side of the interface.

2. Write python code to preprocess data

The screenshot shows a code editor within the EarthDataMiner system. The code is in Python and includes imports for 'os', 'requests', 'urllib', 'numpy', 'pandas', 'matplotlib', 'cv2', 'skimage', 'skimage.filters', 'skimage.measure', 'skimage.segmentation', 'skimage.metrics', 'skimage.registration', 'skimage.feature', 'skimage.util', 'skimage.color', 'skimage.exposure', 'skimage.restoration', 'skimage.metrics', 'skimage.metrics', 'skimage.metrics', 'skimage.metrics'. The code defines a function 'preprocess_data' that takes 'url' and 'region' as input and returns a 'Task' object. An arrow labeled 'Code online, save result as a new file' points to the code. Another arrow labeled 'Preview execution results' points to a small thumbnail of a map showing the results of the preprocessing.

3. Search radar data from Databank develop by CASEarth

The screenshot shows the EarthDataMiner interface with a search form. The search form has a 'Query keyword' field and a 'Query results' table. The table lists search results with columns for 'ID', 'Name', 'Description', and 'Location'. An arrow labeled 'Open DataBank' points to the search form. Another arrow labeled 'Query results' points to the table. The workflow diagram from the first step is also visible in the background.

4. Compute the SDGs Indicator and Visualization

The screenshot shows the EarthDataMiner interface with a code editor and a map. The code editor shows Python code for computing the SDG15.1.1 indicator. An arrow labeled 'Webapp examples' points to the code. Another arrow labeled 'Code for SDG15.1.1' points to the code. The map shows the 'Indicator Tool: SDG15.1.1 Forest cover assessment' with a 'Preview' button. The workflow diagram from the first step is also visible in the background.

Converged Virtual Research Environment : SDG Workbench

➤ Data be accessed by applications/machine Transparently

➤ SDGs Tools

- ✓ Integrated Tools
- ✓ Data Analysis Tools
- ✓ Data Products Tools

➤ Open-source Tools

- ✓ Developing
- ✓ Machine Learning
- ✓ Data Visualization

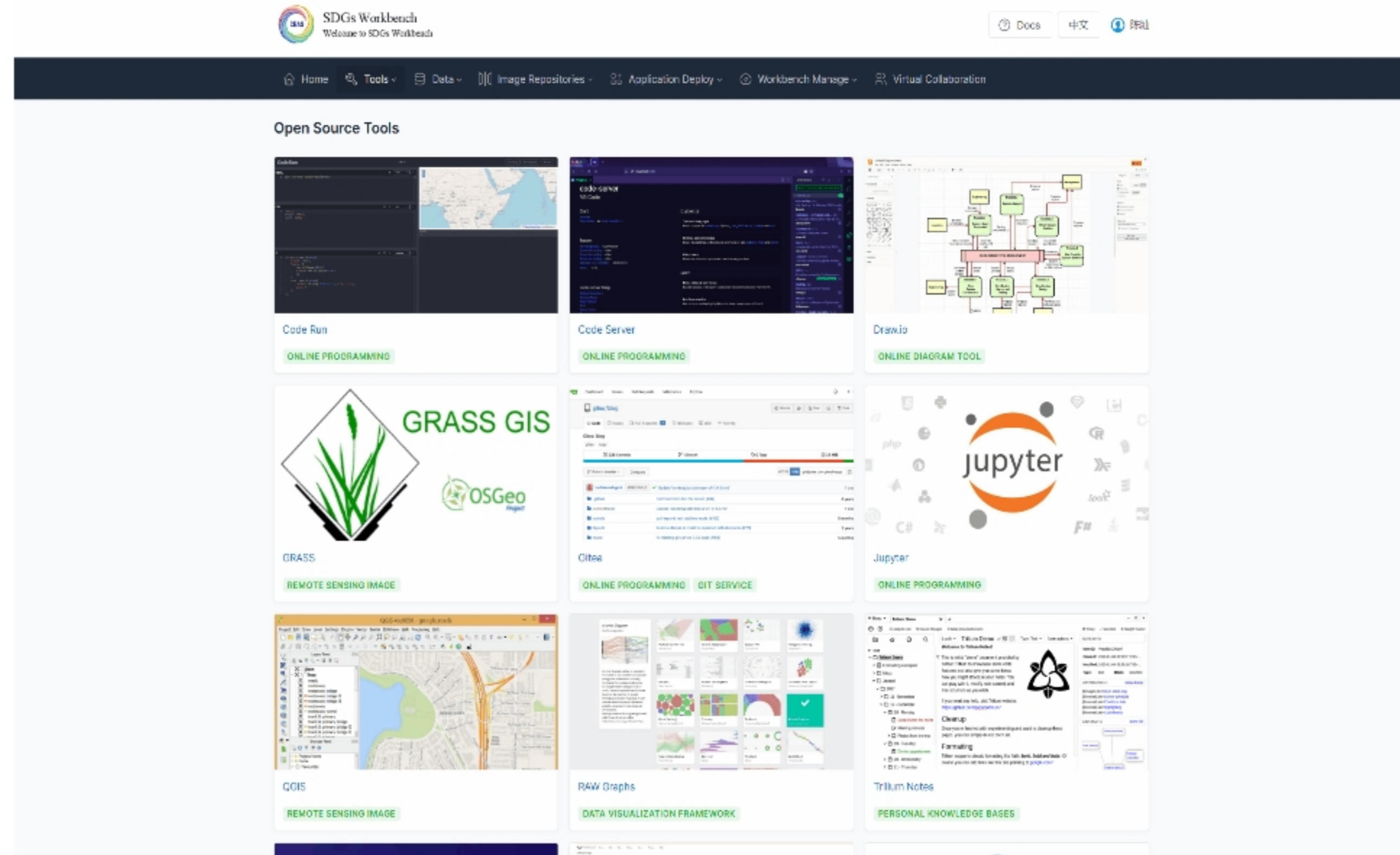
➤ Creating / Using / Releasing Spark

Cluster On-demand

➤ Cloud-Native DevOps CI/CD

➤ Virtual Collaboration

➤ Setup a virtual team or virtual





中国科学院计算机网络信息中心
Computer Network Information Center,
Chinese Academy of Sciences



Thanks for Your Attention

