INTERNATIONAL RESEARCH CENTER OF BIG DATA FOR SUSTAINABLE DEVELOPMENT GOALS
可持续发展大数据国际研究中心

Capacity Building Workshop for Small Island Developing States:Leveraging Big Earth Data to Evaluate the SDGs Progress

AI Powered Big Earth Data Platform form SDGs: Infrasturcture, System and Applications

Yuchu Qin (qinyc@aircas.ac.cn)

3 Sep 2025

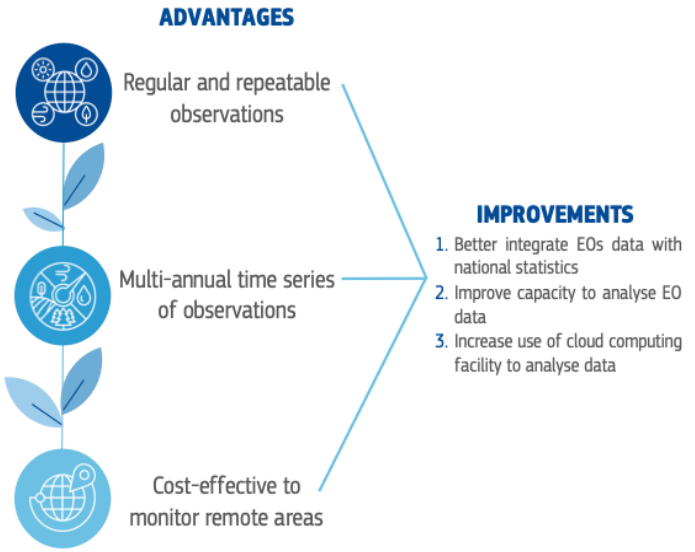
CONTENT

CONTENT

- AI-Big Data for SDGs with Cloud Computing
- Infrastructure at CBAS
- Big Earth Data Platform
- Applications for SDGs
- Lab Exercise

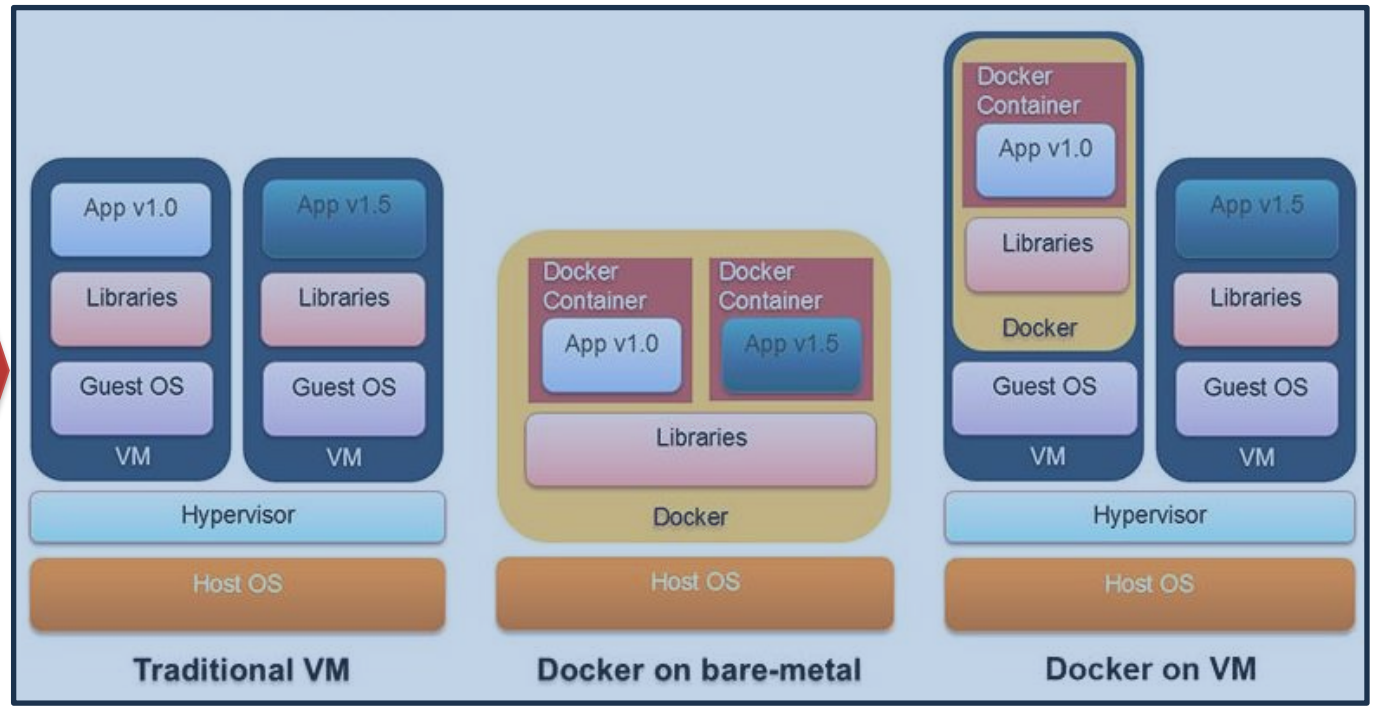
- UN introduced SDGs in 2015, to be completed by the year 2030
- 17 Sustainable Development Goals (SDGs)
- 169 SDGs targets
- 231 unique indicators

Earth Observation for SDG Monitoring



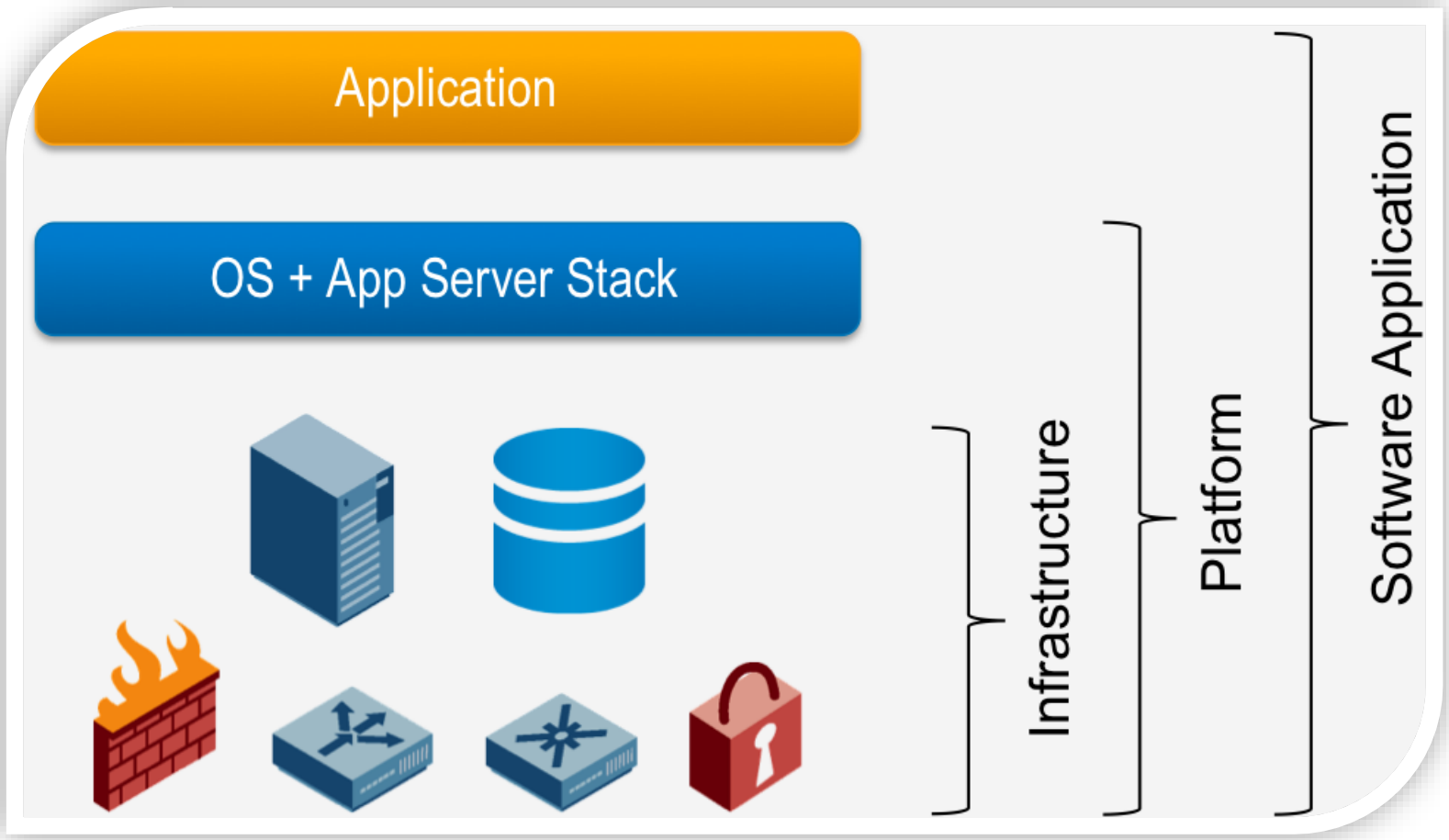
Source: https://eo4society.esa.int/wp-content/uploads/2021/01/EO_Compendium-for-SDGs.pdf

Cloud Computing



From Hardware to software-defined resources: VIRTUALIZATION

Cloud Computing



Google 地图 谷歌 中国

Google 日历

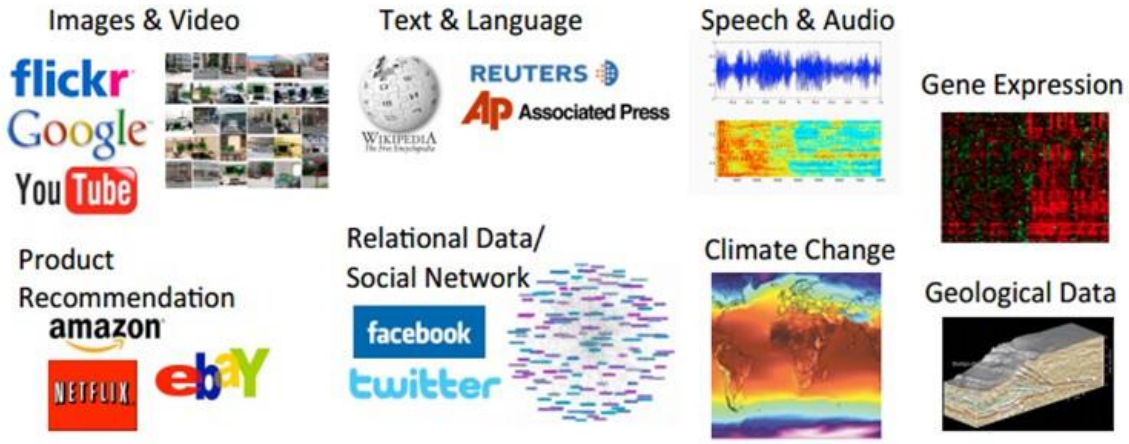
Gmail

Google Earth Engine

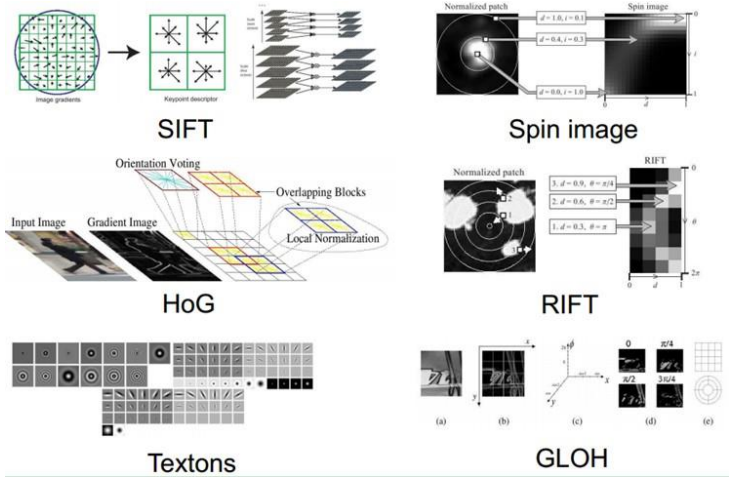
Google app engine

On-demand scalable distributed environment for computing, storage and networking

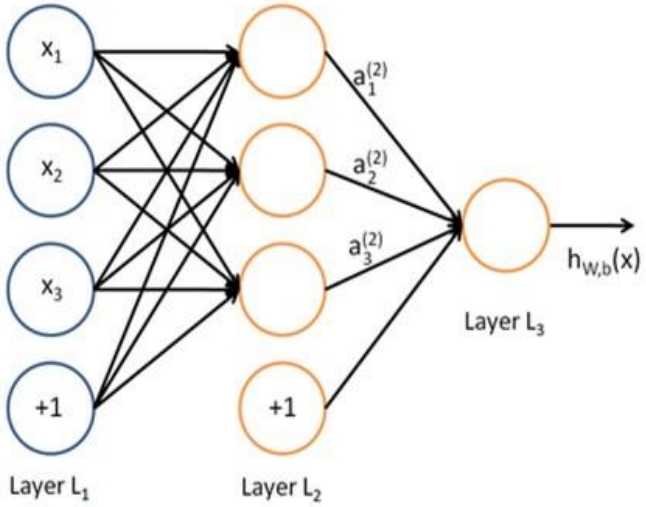
ABC for AI

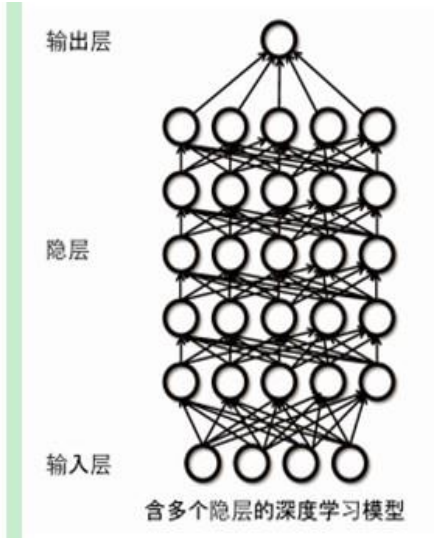


Conventional solution for machine learning:

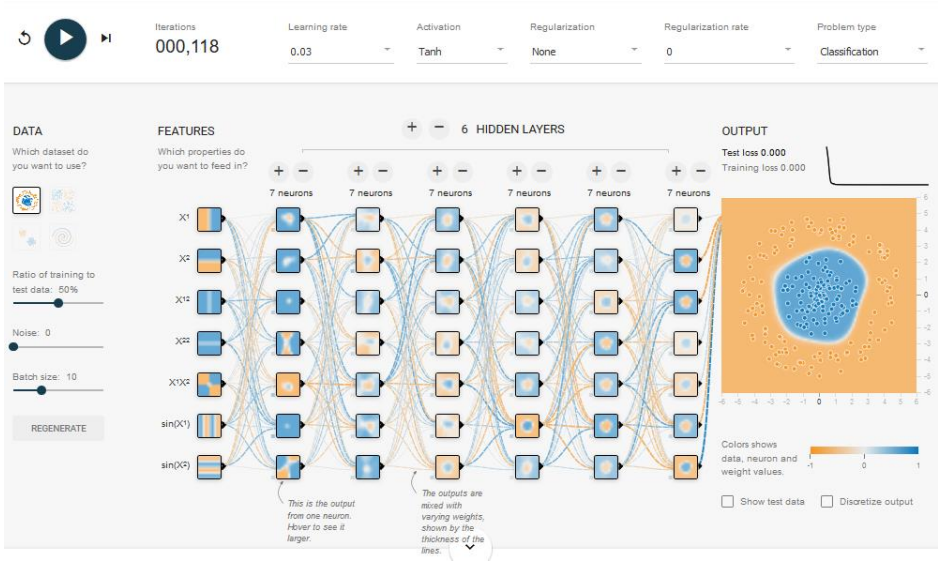


ABC for AI



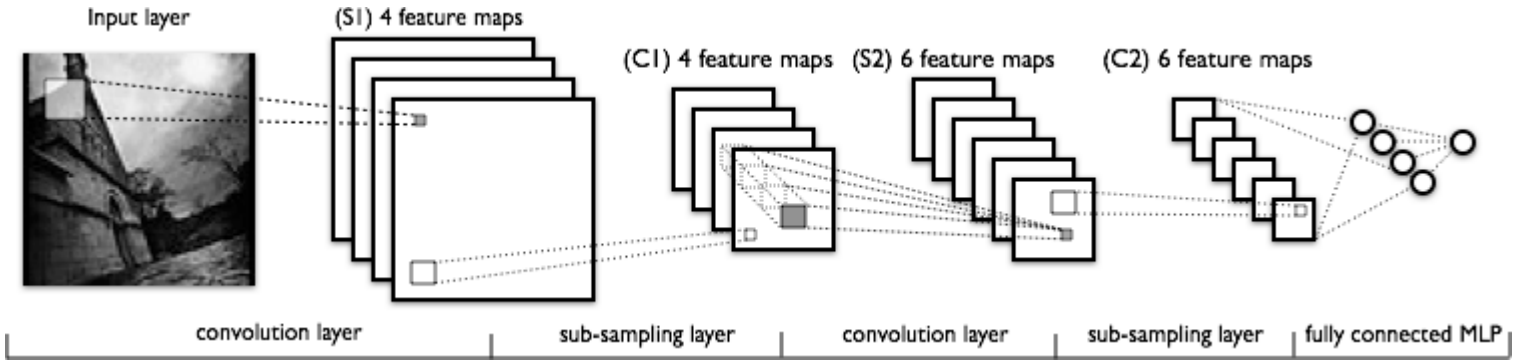


Neural Network: From shallow to deep



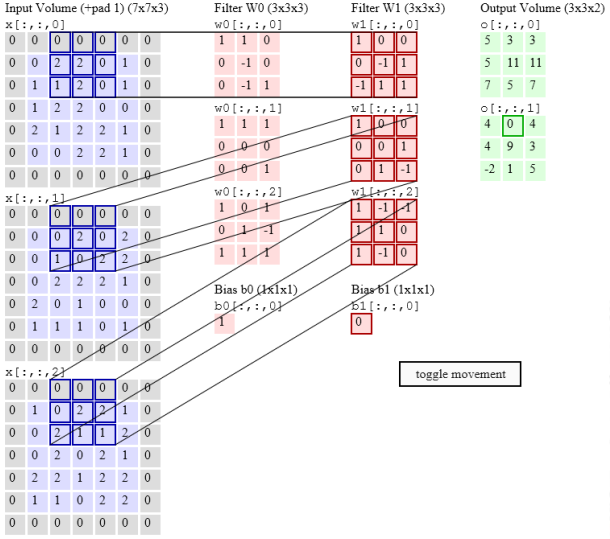
DNN: Google Tensorflow playground

ABC for AI

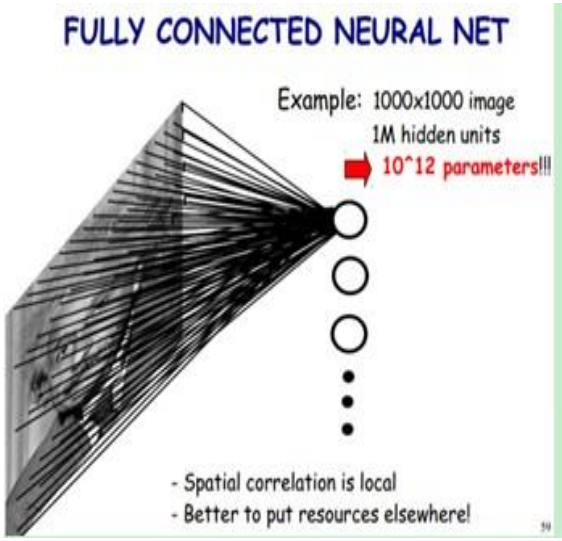


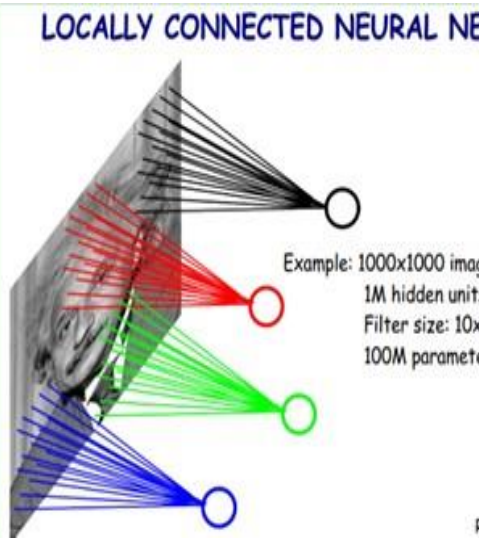
Convolutional Neural Network

CNN: Convolutional Operator for Image

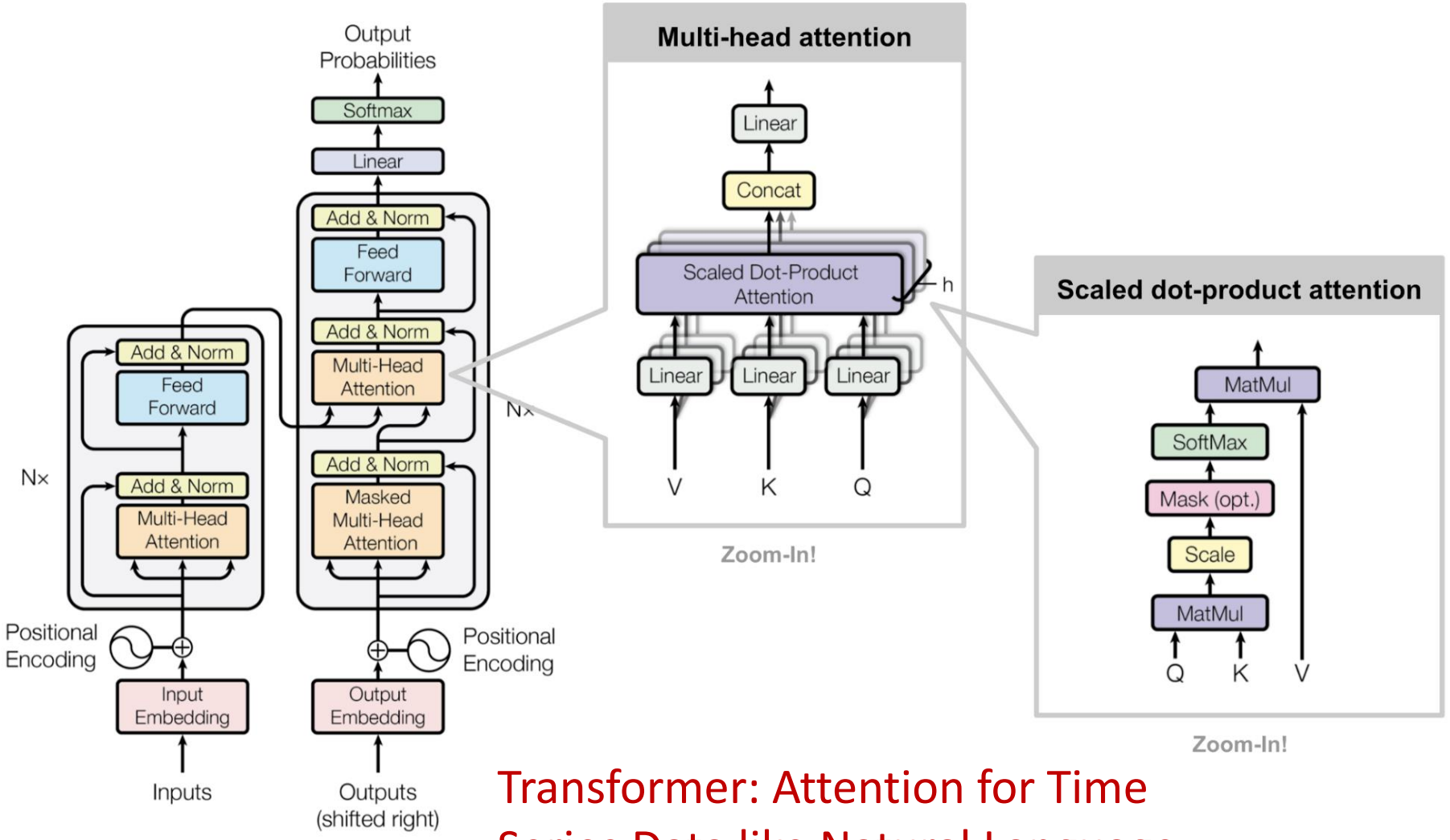


Convolutional filter





ABC for AI

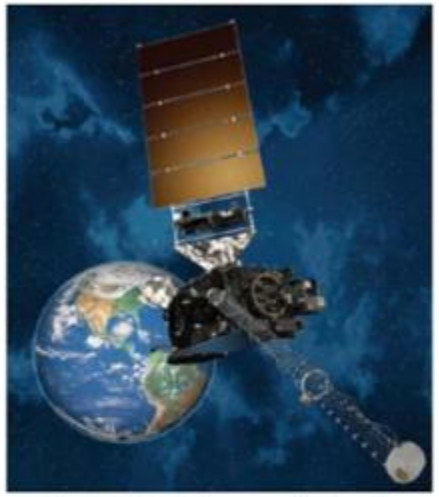


Transformer: Attention for Time Series Data like Natural Language

Data is the key source of SDGs

Ground-based instruments used to observe precipitation include rain gauge tipping buckets, cylinders, and disdrometers & radar systems [top]

A sensor pod from NASA – Jet Propulsion Laboratory



The GOES-R Series—a collaborative program between NASA and NOAA.

Argo floats are used to observe the ocean [image from Commonwealth Scientific and Industrial Research Organization]

Researchers with the University of Alaska-Fairbanks (UAF) use small aircraft such as the Havilland DHC-3 Otter. Credit: UAF

Space-based
Satellites

Airborne

Ground-based

In Situ

The data GAP of SDGs

可持续发展目标

Basic information about global SDGs

- 17target，230+indicators
- UN TFM，Supporting SDGs realization via STI

- **Tier I**：clear method，has data （33%）
- **Tier II**：clear method, no data （**30%**）
- **Tier III**：no method，no data （**37%**）

6 SDGs assessment in CBAS

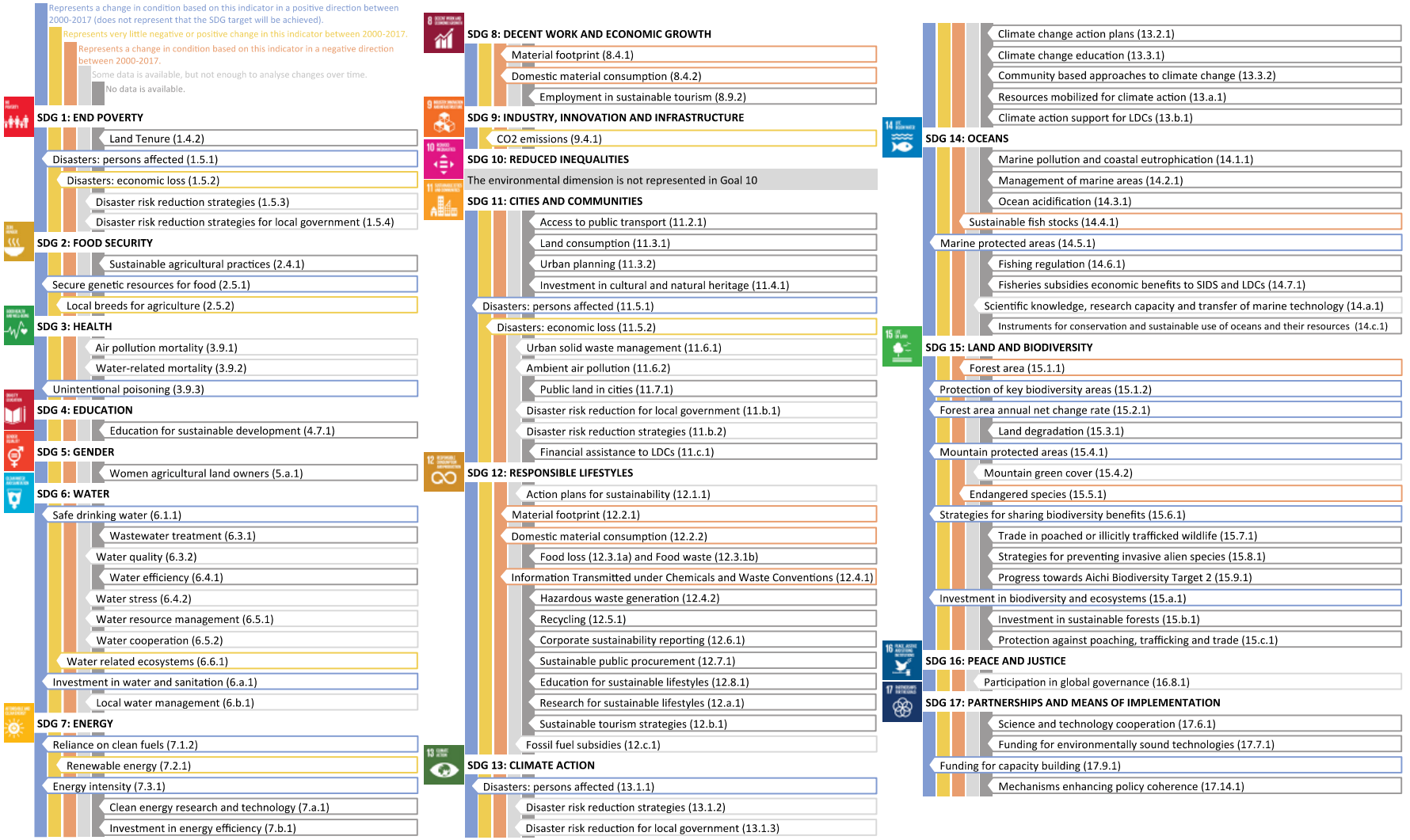
THE WORLD BANK IBRD • IDA | WORLD BANK GROUP　UNISDR

FAO　UN environment

More than 67 international organizations carry out the SDGs indicator assessment

Of the 93 environmental related SDGs indicators, there are 22 (23 percent) with good progress, not sufficient data to assess progress (68 percent) or unlikely realize without upscaling action (9 percent).

Figure 1. Global Scorecard on the environmental dimension of the SDGs based on extrapolation of data to 2030



The Paradigm of AI+Data for SDGs

OPEN Data + OPEN Models -> Open Science for SDGs

MOODY'S ANALYTICS

WIKIPEDIA The Free Encyclopedia

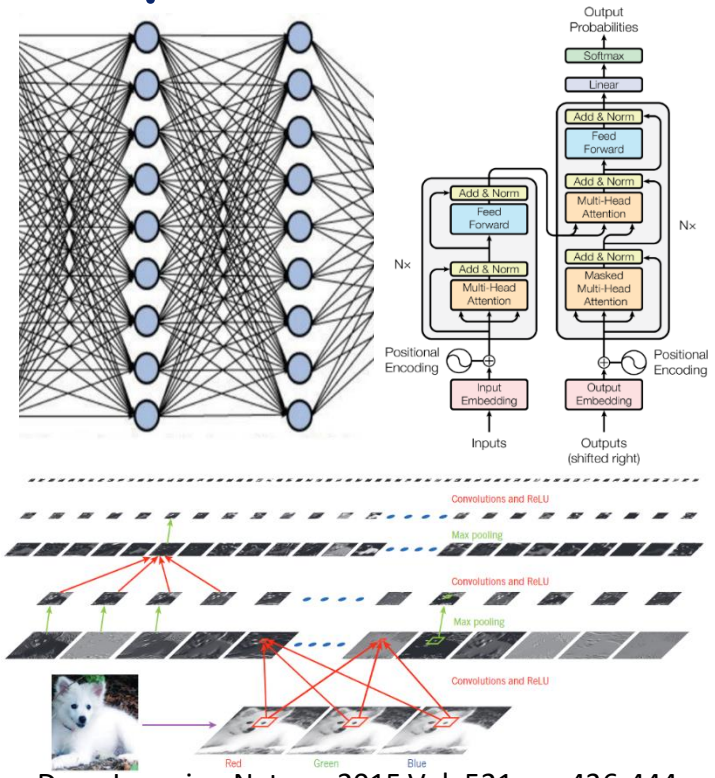
THE WORLD BANK IBRD • IDA

UNdata A world of information

Data

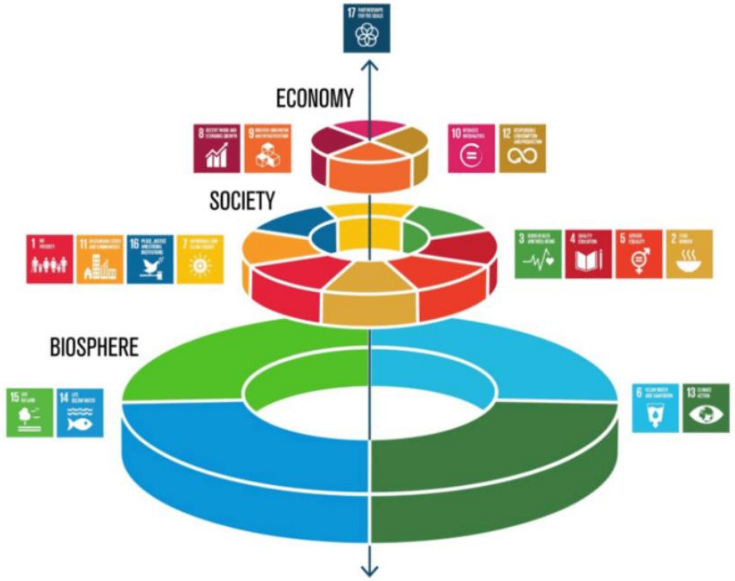
\+

Deep Neural Networks



Deep Learning,Nature, 2015,Vol. 521, pp 436-444

Models



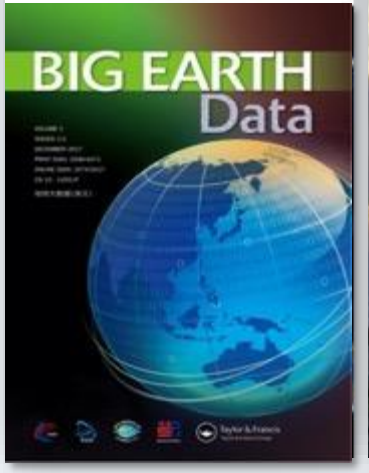
http://dx.doi.org/10.5751/ES-08748-210341

SDG indicators

CONTENT

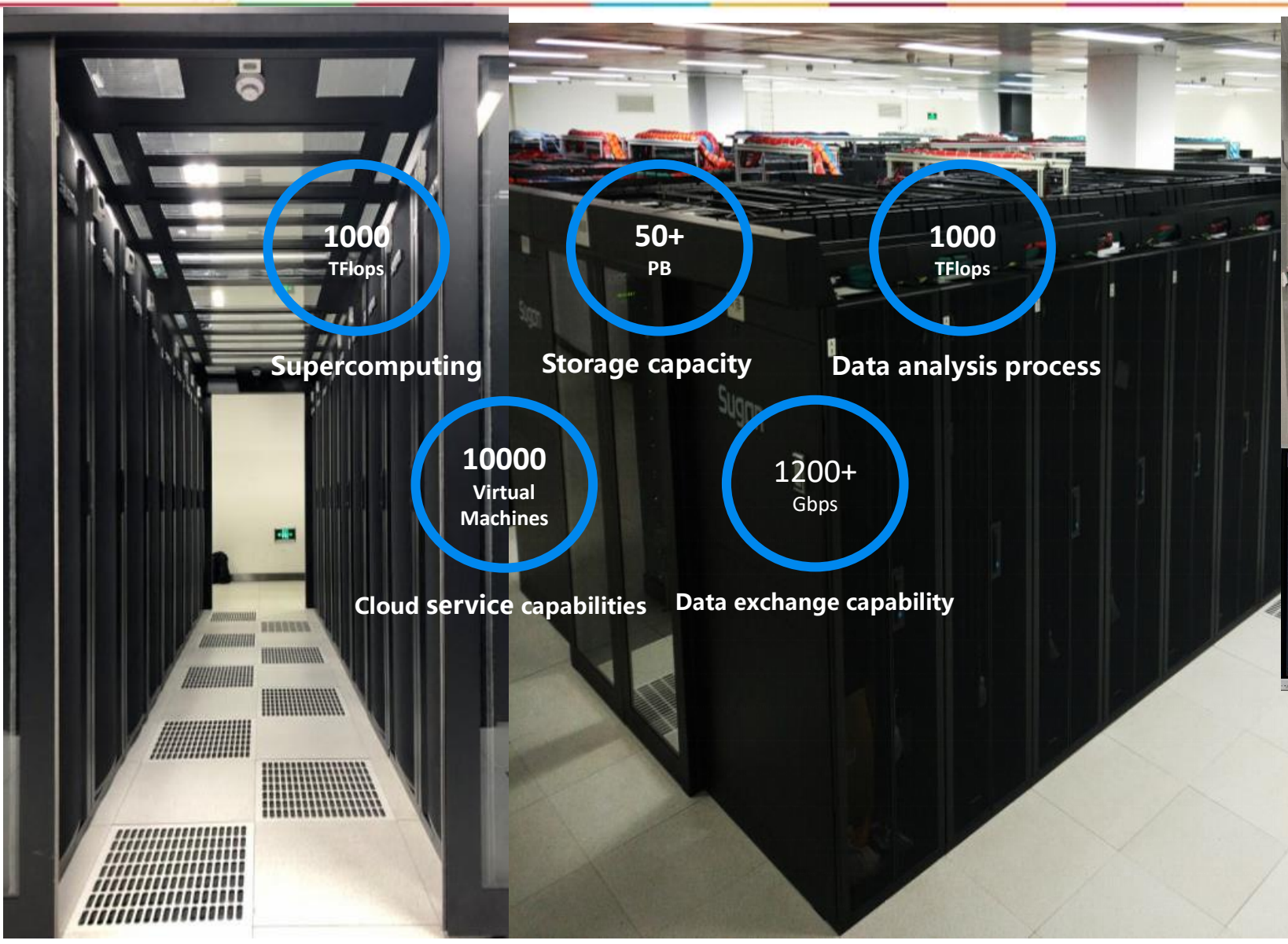
- AI-Big Data for SDGs with Cloud Computing
- Infrastructure at CBAS
- Big Earth Data Platform
- Applications for SDGs
- Lab Exercise

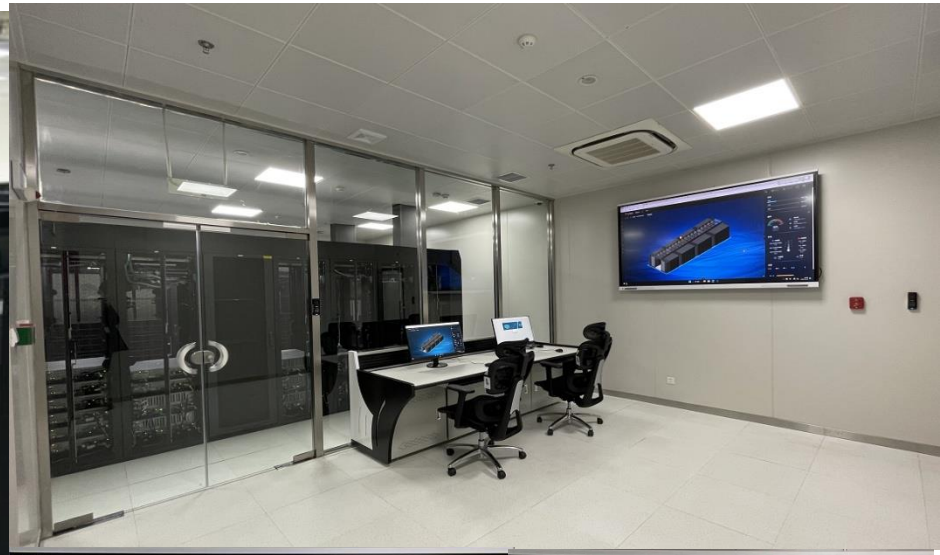
Mission of CBAS

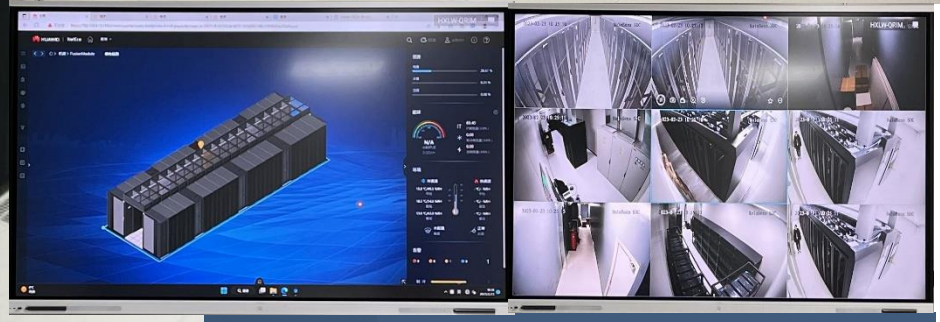


- Develop SDG **data infrastructure and information and data products**
- Provide new **knowledge** for SDG monitoring and evaluations
- Develop and launch **a series of SDG Satellites**
- Establish a **think tank** for STI to promote SDGs
- **Capacity development** for SDGs in developing countries

SDG Cloud Platform







- Cloud enabled
- GPU cluster for AI
- with VMs and Containers

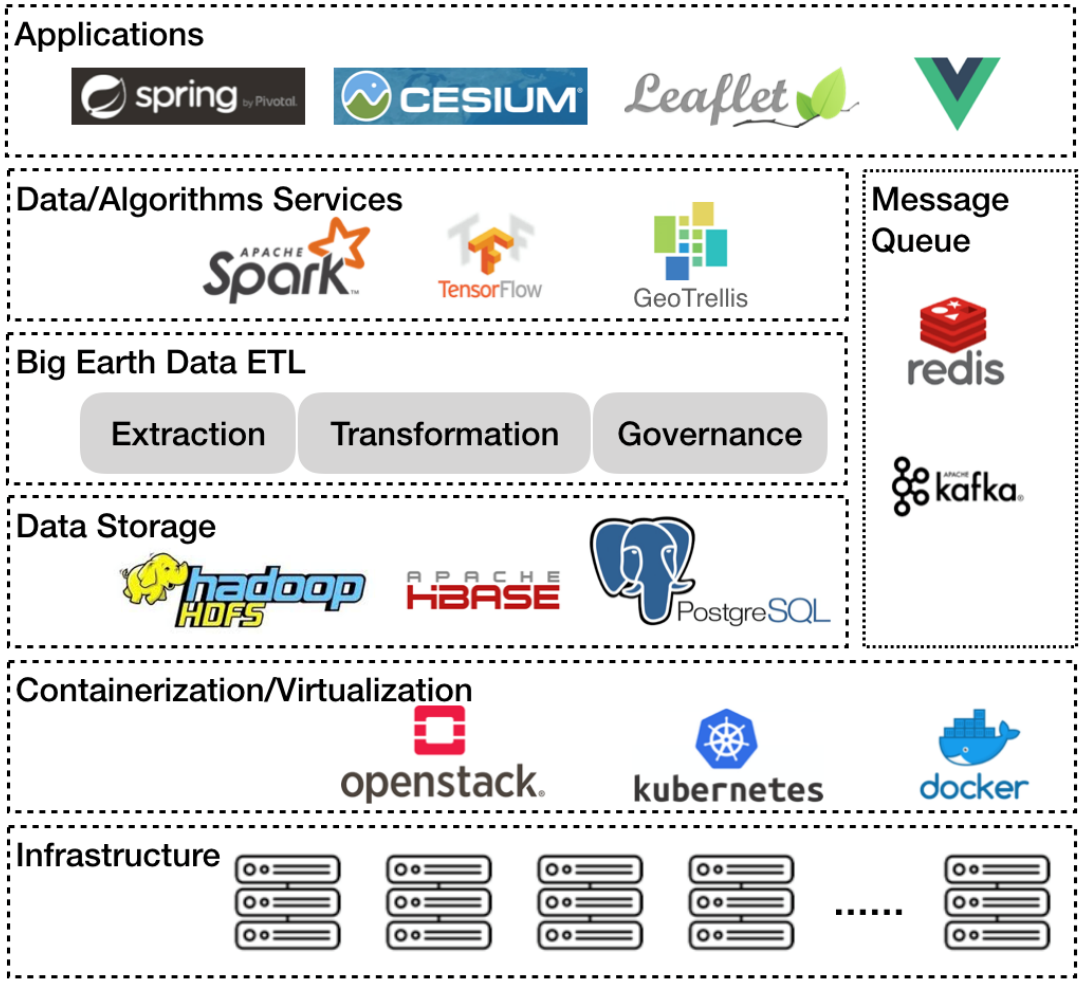
The Platform: Tech Goals

- **Manage** huge amount of Big Earth Data at scale
  - Multi-source data integration(ETL)
  - Multi-dimentional modelling
  - Data quality control and metadata compiling
  - Data access in distributed enviroment with APIs

- **Process** Big Earth Data at scale
  - Image processing algorithms at cloud computing platform
  - Feature data processing algorithms
  - Machine learning(Deep learning)
  - Algorithm access with restful APIs

- Web-based **application** system
  - User session control
  - Report generation

Architecture: Tech Stacks



Tech Stacks

Web-based Visuliazation

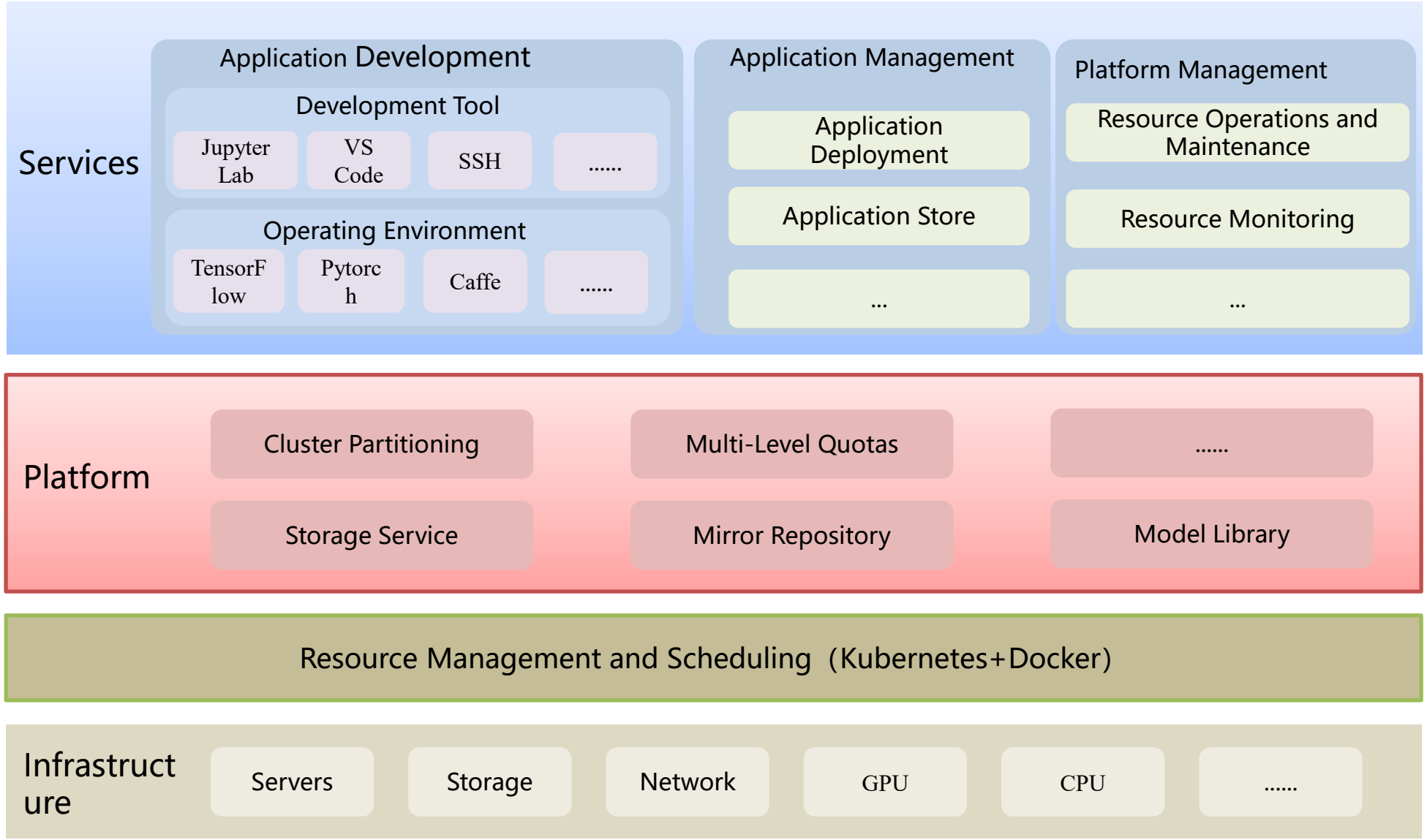
Big Data-driven Analytics

Big Earth Data Management

Scalable Enviroment

Cloud Enabled

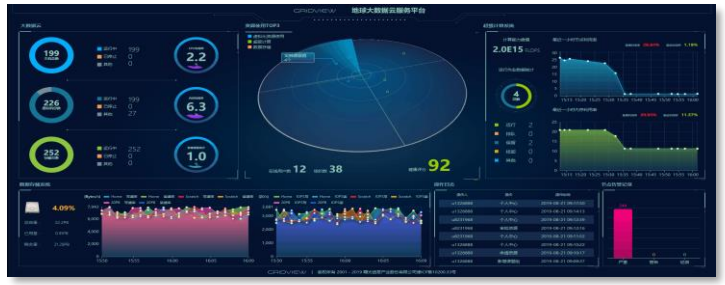
Architecture: Tech Stacks



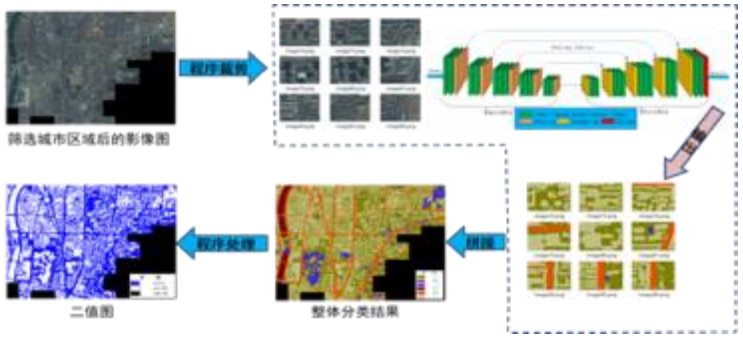
- Containerized Microservices Architecture
- Cloud Computing and Storage Resources
- Online flexible development environment
- Multi-user resource sharing
- From IAAS\PAAS to SAAS

Cloud Services for SDGs

- **HPC & Cloud Computing**
  - 12000 cores for meteorological and oceanographic modelling
  - 10000 VMs for high-throughput computing



- **Data Storage, Publishing and Sharing**
  - 50PB storage capacity
  - Data linking, data recommendation and sorting
  - Data discovery in various ways



- **Big Data Mining & Analysis in Cloud**
  - DataBank: a complete chain for multidisciplinary data, computing and analysis
  - DataBox: an efficient query & access engine for RS data

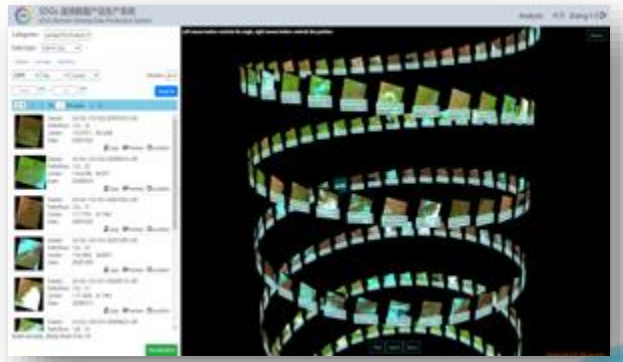
- **Application Service**
  - Online Computing and Processing Toolkit for Specific SDG indicators
  - On-demand Data Processing and Analysis

CONTENT

- AI-Big Data for SDGs with Cloud Computing
- Infrastructure at CBAS
- Big Earth Data Platform
- Applications for SDGs
- Lab Exercise

SDG Big Data Platform

- **The *SDG Big Data Platform*** provides rich data resources and cloud services for SDGs monitoring and assessment
  - With **16PB 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.



Remote Sensing Data Engin

SDGs Data Products

SDGs workbench

SDGs assessment tools

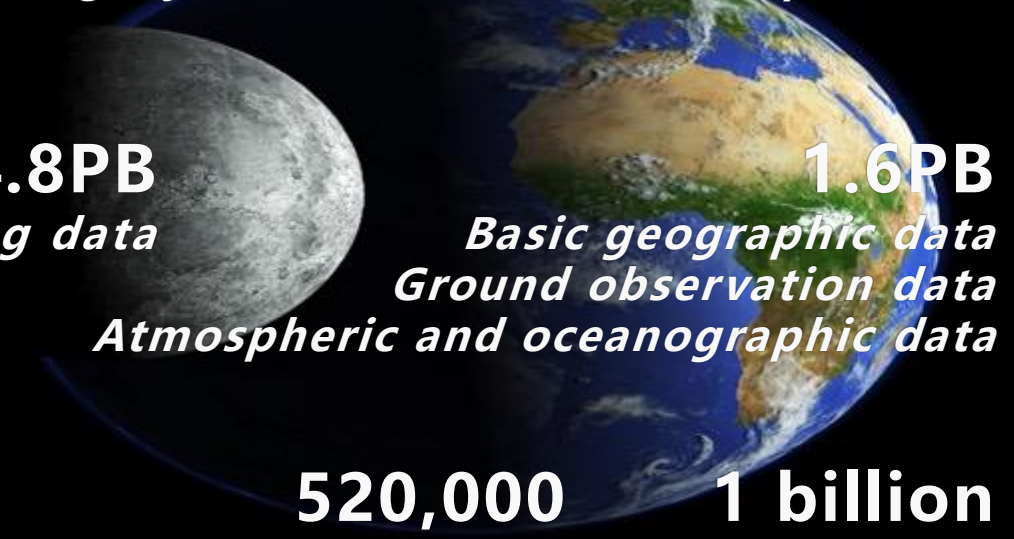
Http://sdg.casearth.cn

Collected data resources for SDGs

16.4PB
Total Data

40 years
Satellite imagery

7.7 million scenes
Satellite data products



10PB
Bio-ecological data

4.8PB
Remote sensing data

1.6PB
Basic geographic data
Ground observation data
Atmospheric and oceanographic data

920,000
GBDB data records

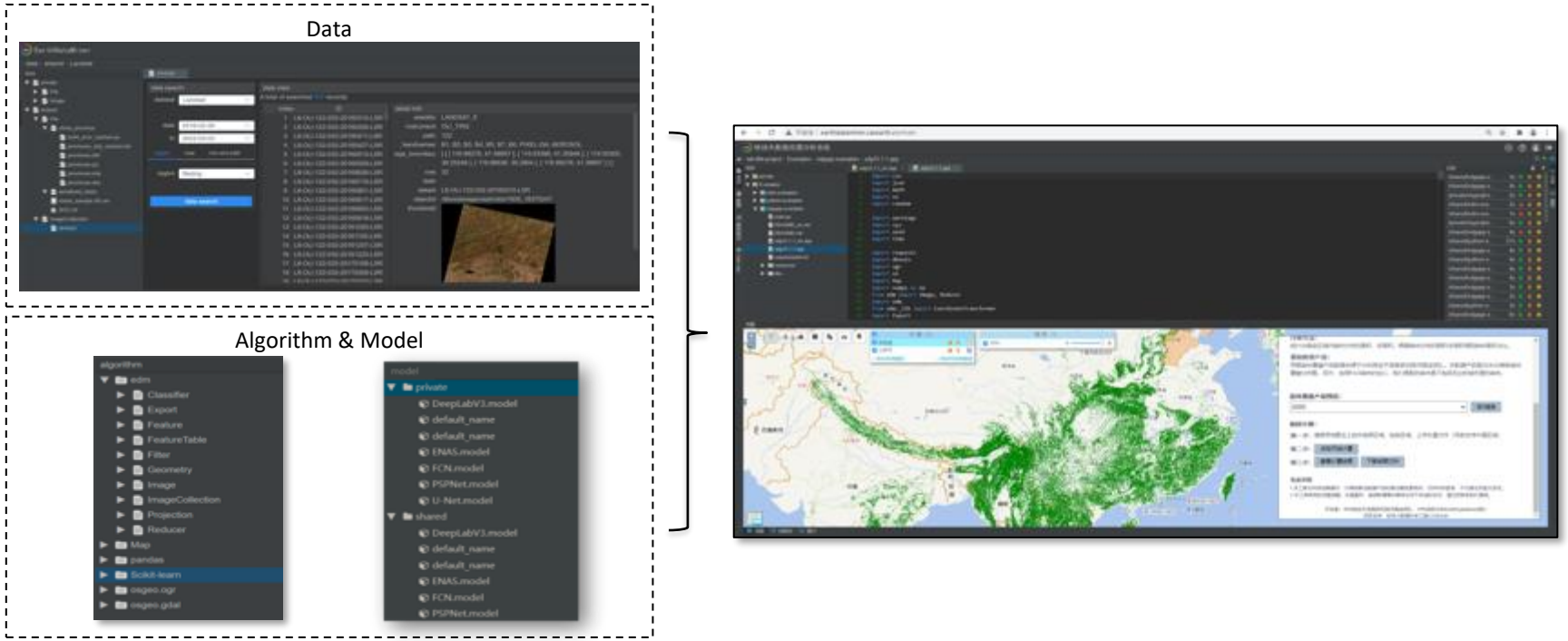
4.08 million
Lists of biological species in China

520,000
Microbial resource data

1 billion
Omics data

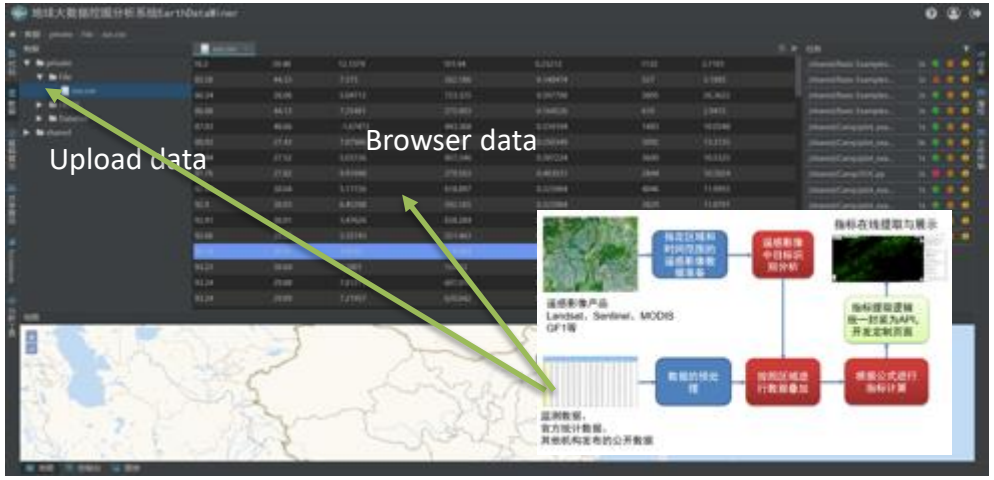
Data Analysis As Service: EarthDataMiner

- Online interactive data analysis environment
  - Integrated with DataBank: upload models, select data, and process data products through instruction operations
  - Algorithm & Model library: more than 150 algorithms developed and provide cloud service: FAAS(Function As A Service)
  - Web IDE: supporting users to write data analysis code (Python) online

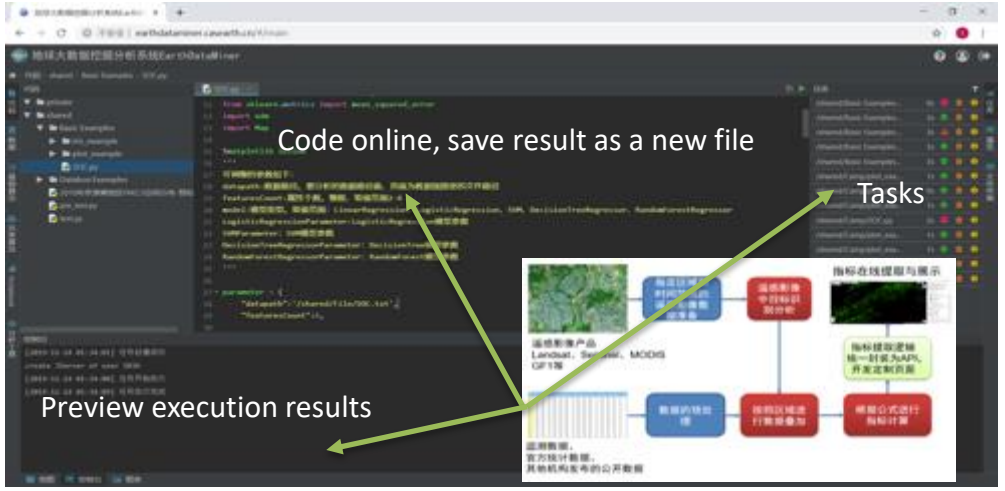


SDGs indicator on-demand computing

1. Upload data to the EarthDataMiner system



2. Write python code to preprocess data



3. Search radar data from Databank develop by CASEarth

4. Compute the SDGs Indicator and Visualizaiotn

SDGs indicators on demand analysis Tools

SDG15.3.1 Land degradation index assessment

SDG6.6.1 Surface water change over time indicator

SDG13.2.2 Annual average CO2 concentration evaluation

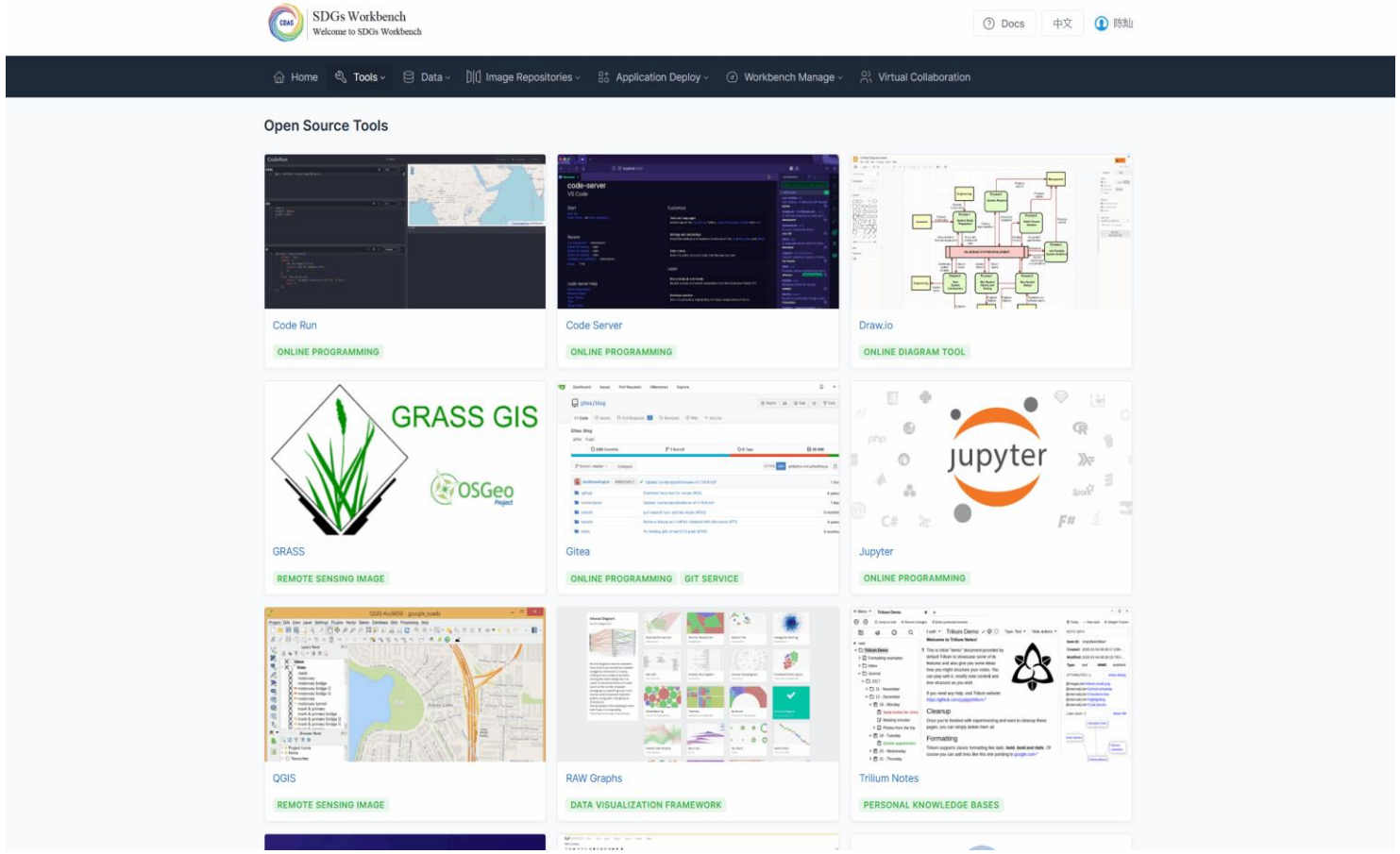
SDG13.1.1 Natural disaster impact evaluation

SDG15.1.1 Forest coverage detection

SDG11.3.1 Global urbanization index monitoring

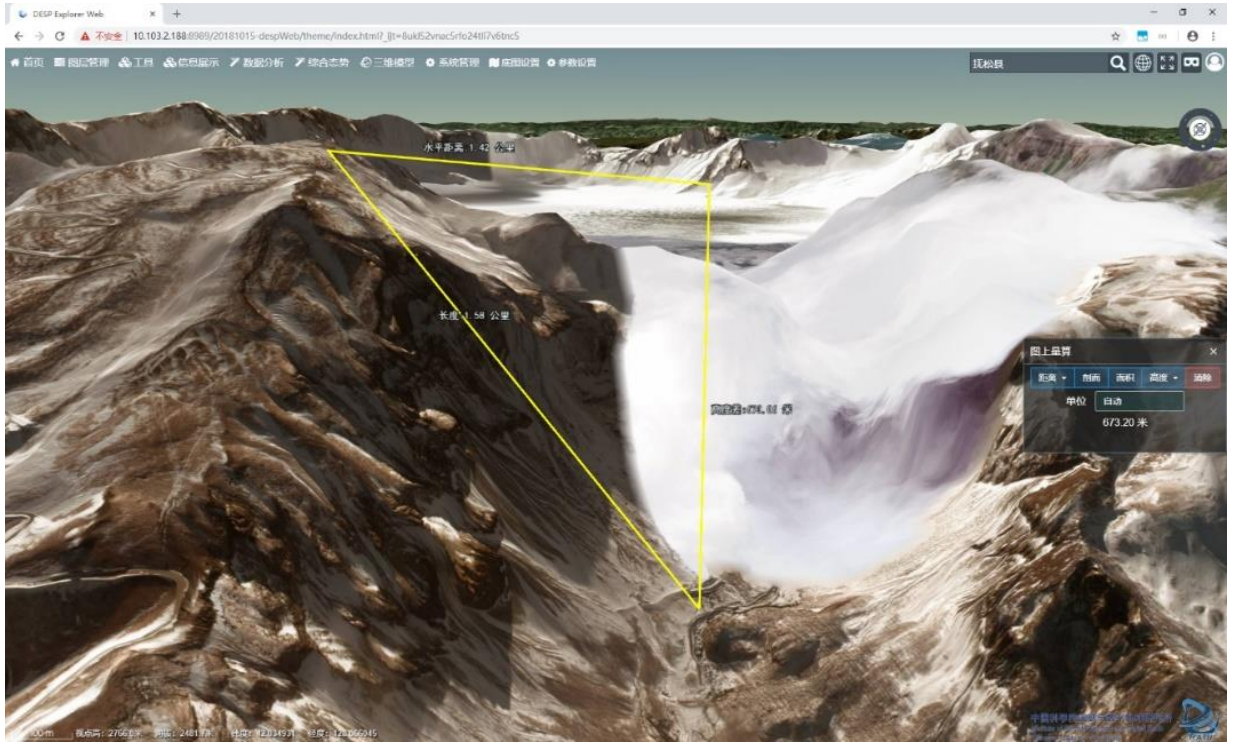
Converged Virtual Research Environment : SDG Workbench

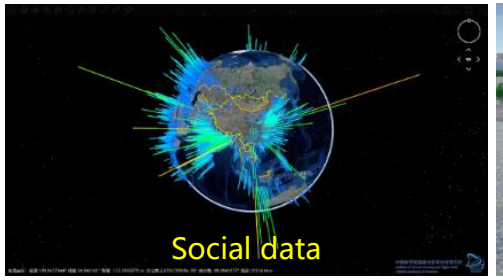
- Data be accessed by applications/machine Transparently
- SDGs Tools
  - ✓ Integrated Tools
  - ✓ Data Analysis Tools
  - ✓ Data Products Tools
- Open-source Tools
  - ✓ Developing
  - ✓ Machine Learning
  - ✓ Data Visualization
- Creating / Using / Releasing Spark Cluster On-demand
- Cloud-Native DevOps CI/CD
- Virtual Collaboration
  - Setup a virtual team or virtual organization



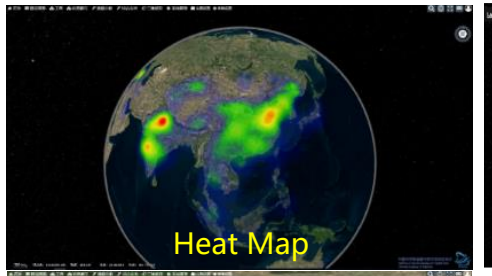
DESP: digital earth foundation platform (DESP Explore)

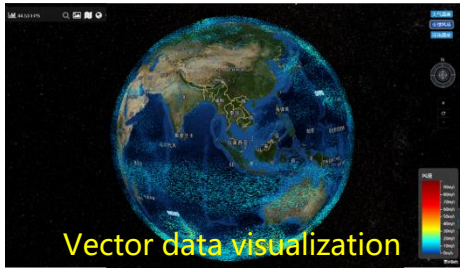
- **Extensibility/flexibility: Plugin-framework**
- **C/S and B/S : Uniform interface style**
- **DESP Development Engine (kit)**

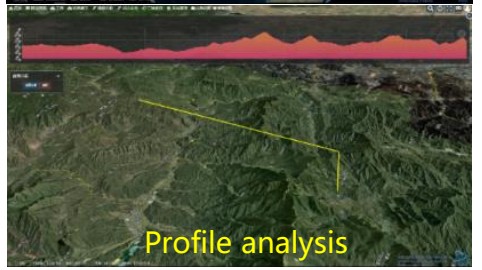




C/S Version







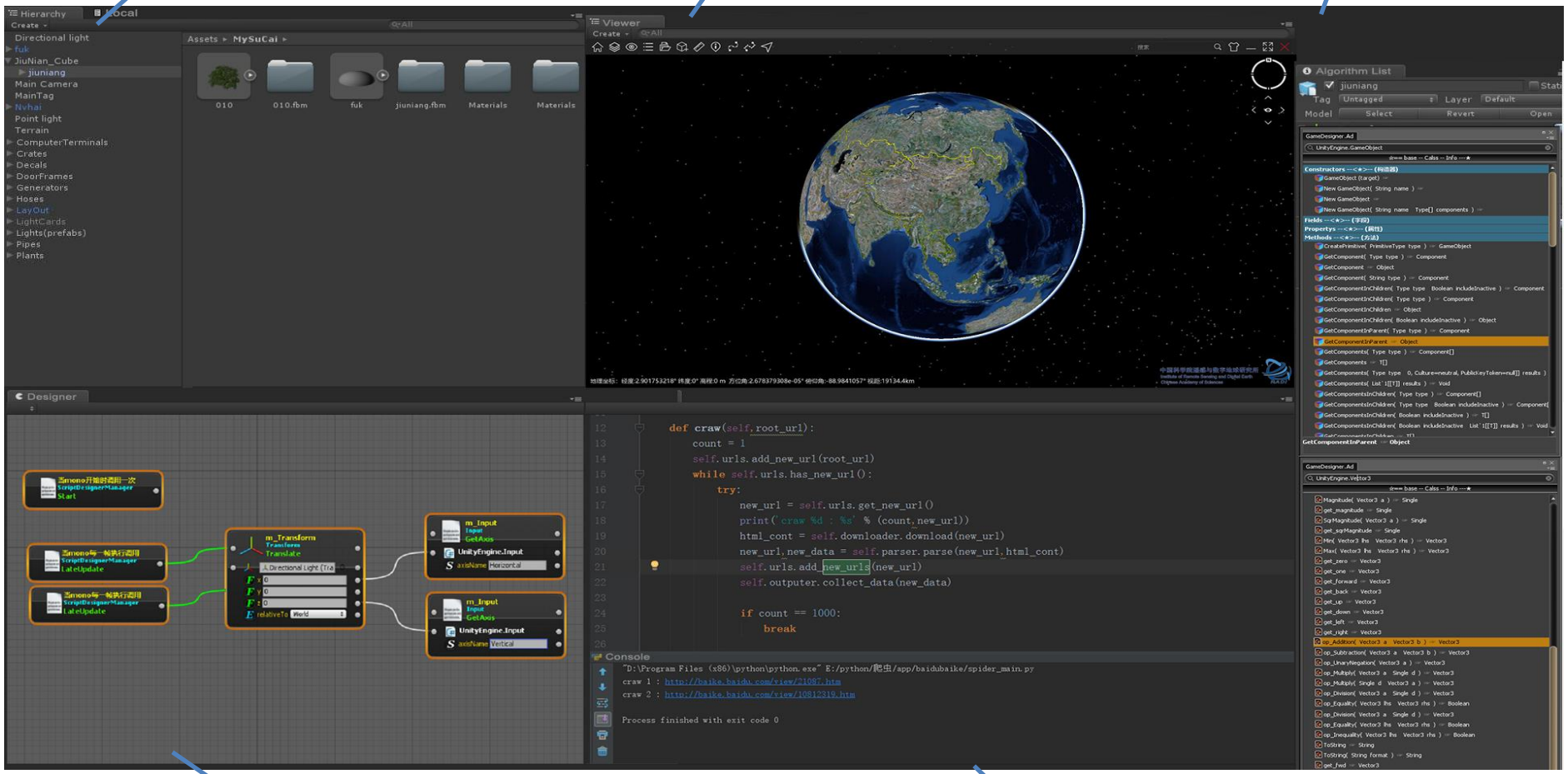
B/S Version

DESP: digital earth foundation platform (DESP Explore)

Resources (data)

2D/3D View

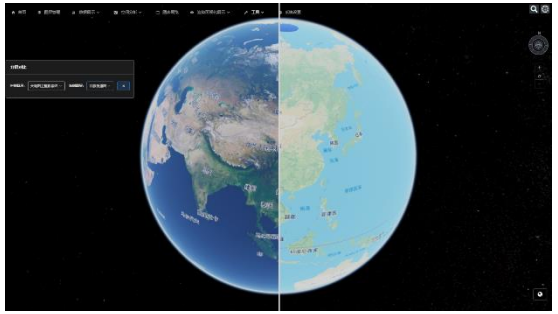
Algorithms

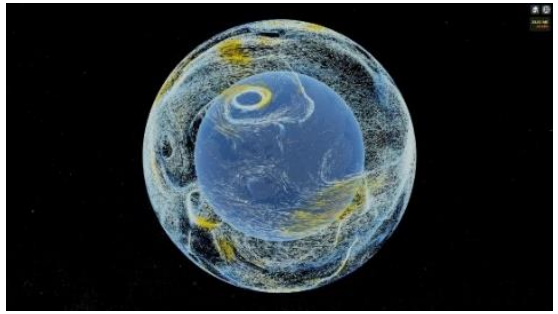


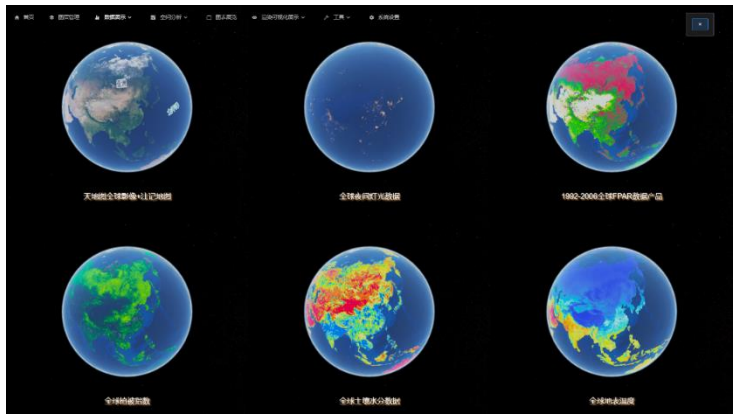
Model Editor

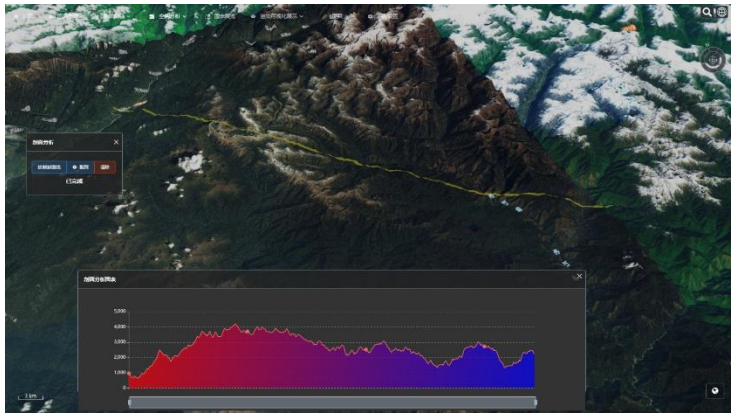
Code Editor

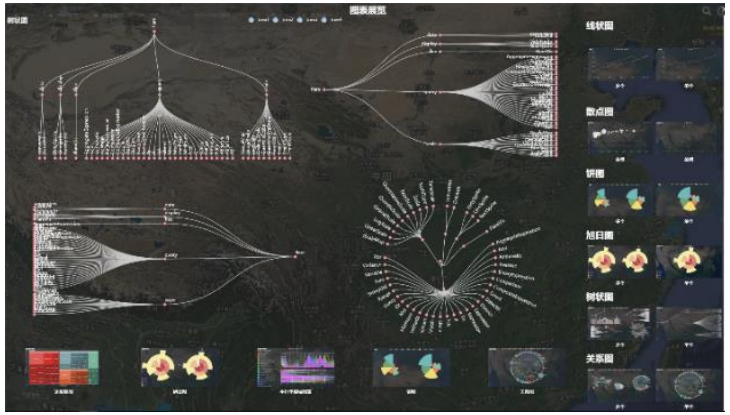
DESP: digital earth foundation platform (DESP Explore)





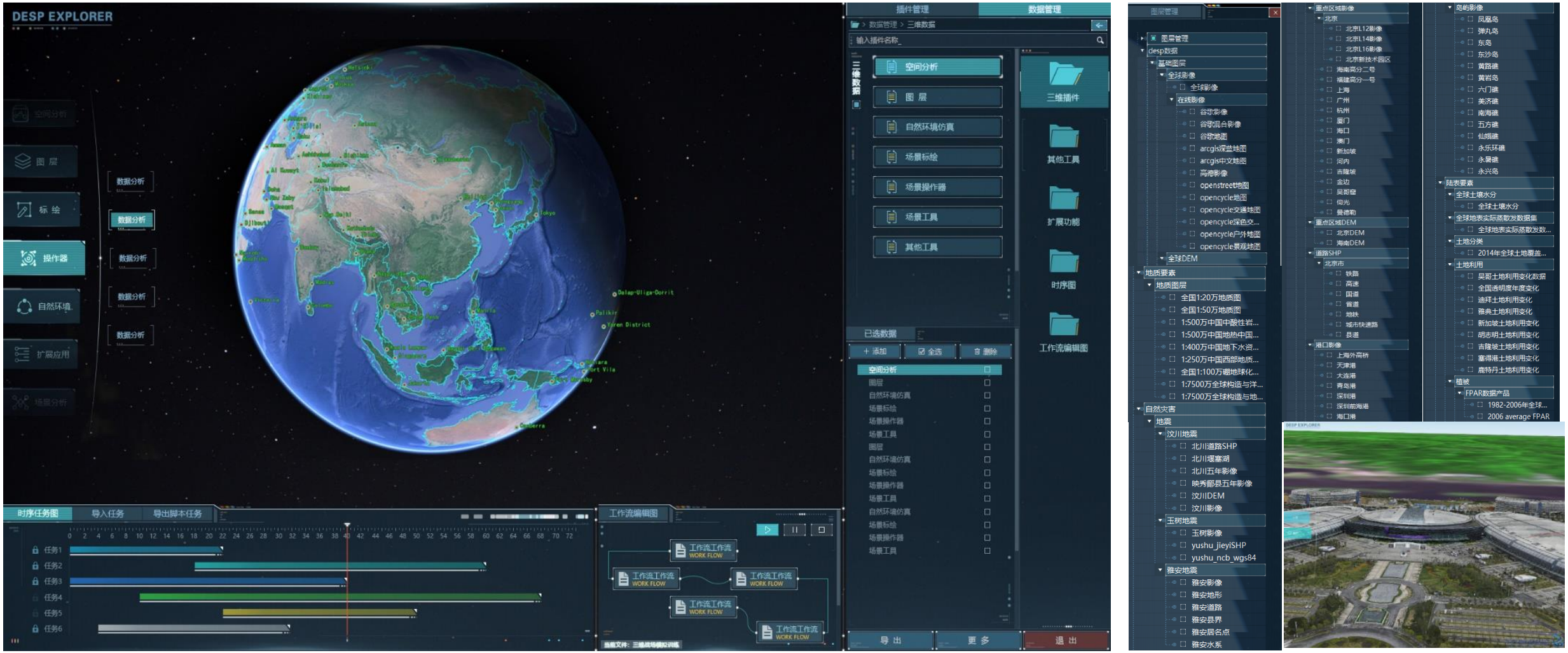






DESP: digital earth foundation platform (DESP Explore)

Desktop Version



CONTENT

- AI-Big Data for SDGs with Cloud Computing
- Infrastructure at CBAS
- Big Earth Data Platform
- Applications for SDGs
- Lab Exercise

Visualizing SDGs progress at scale

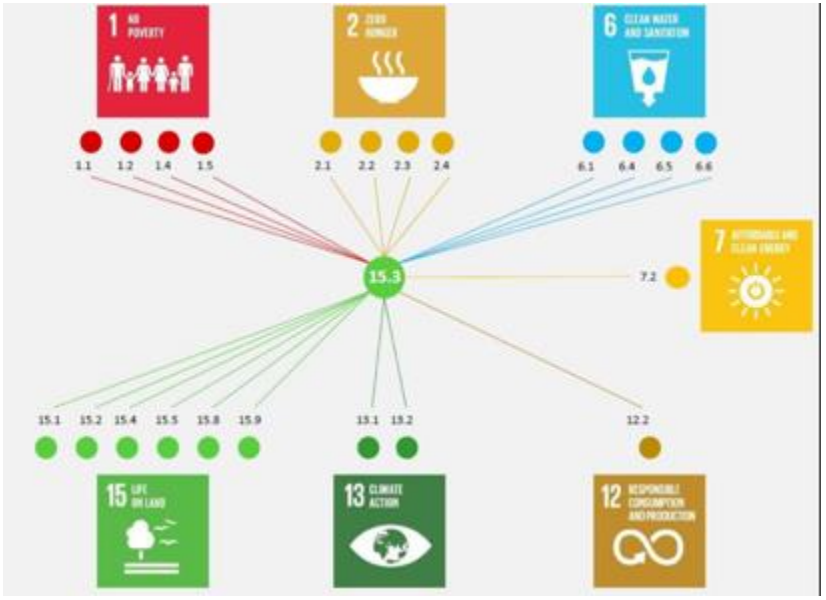
Applications for the Public and Decision-Makers

- **Integration of Multiple Data Sources：United Nations Statistics Division, World Bank, World Health Organization, Wikipedia**
- **Visualizing indicators for the 17 SDGs at National scale**
- **Multi-dimensional analysis**

National-level assessment of SDGs indicators and Data Visualization

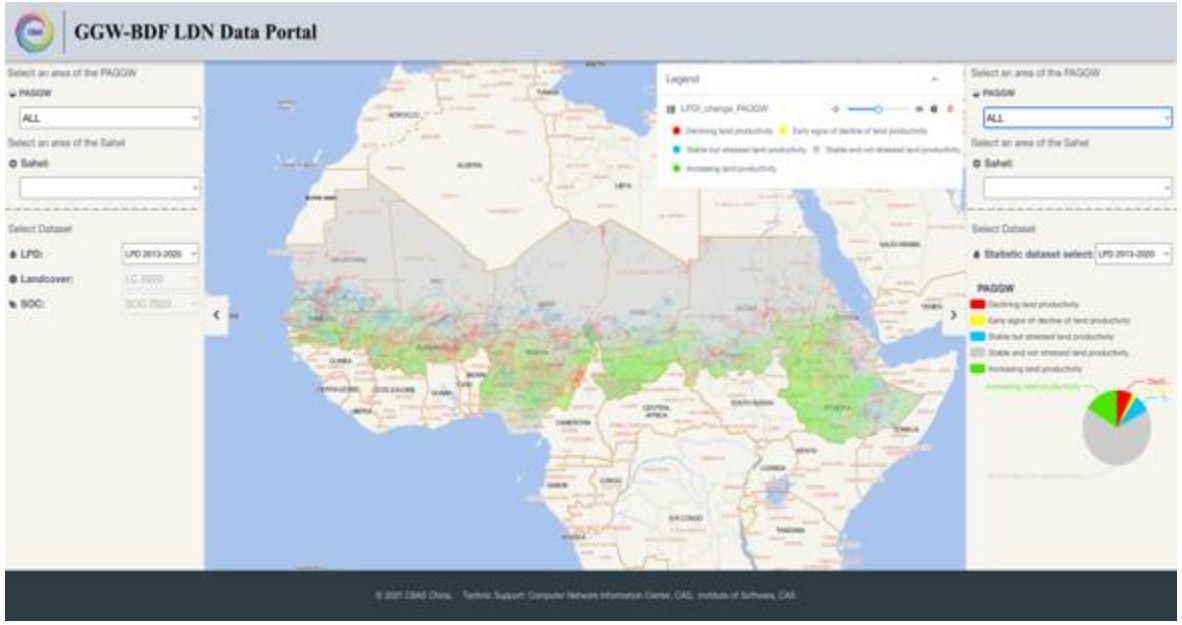
SDG15.3.1 LDN

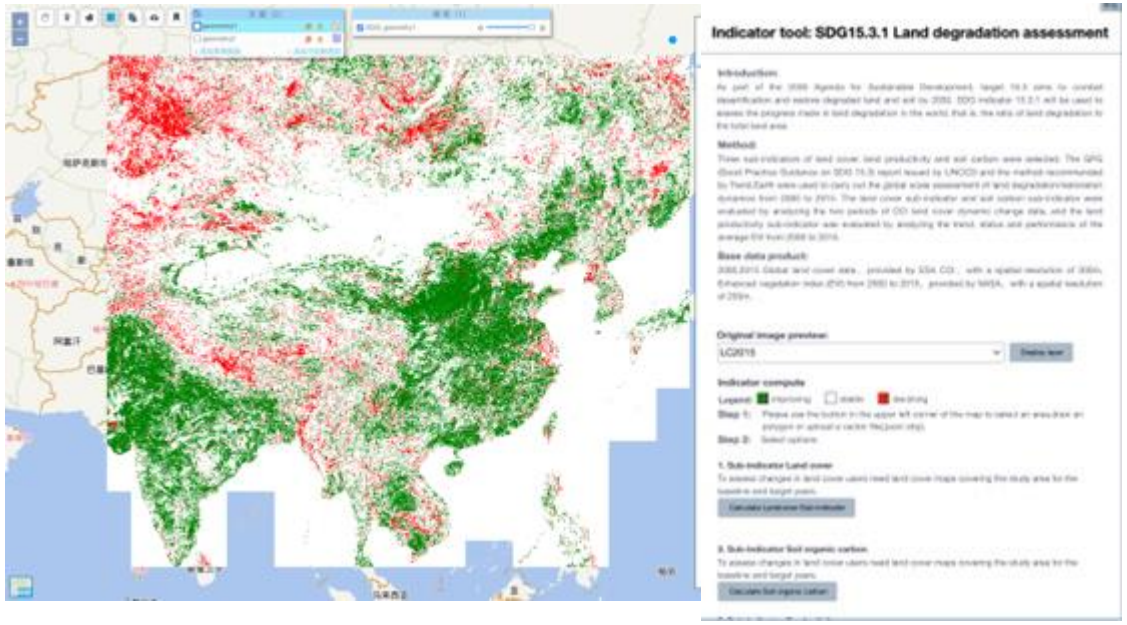
- Goal 15, Target 3
  - *By 2030, combat desertification, restore degraded land and soil, including land affected by desertification, drought and floods, and strive to achieve a land degradation-neutral world*
  - *UNCCD: custodian agency (← CBD, UNFCCC, UNEP)*
- SDG indicator 15.3.1:
  - *Proportion of land that is degraded over total land area*



GGW-BDF & SDG15.3.1 LDN

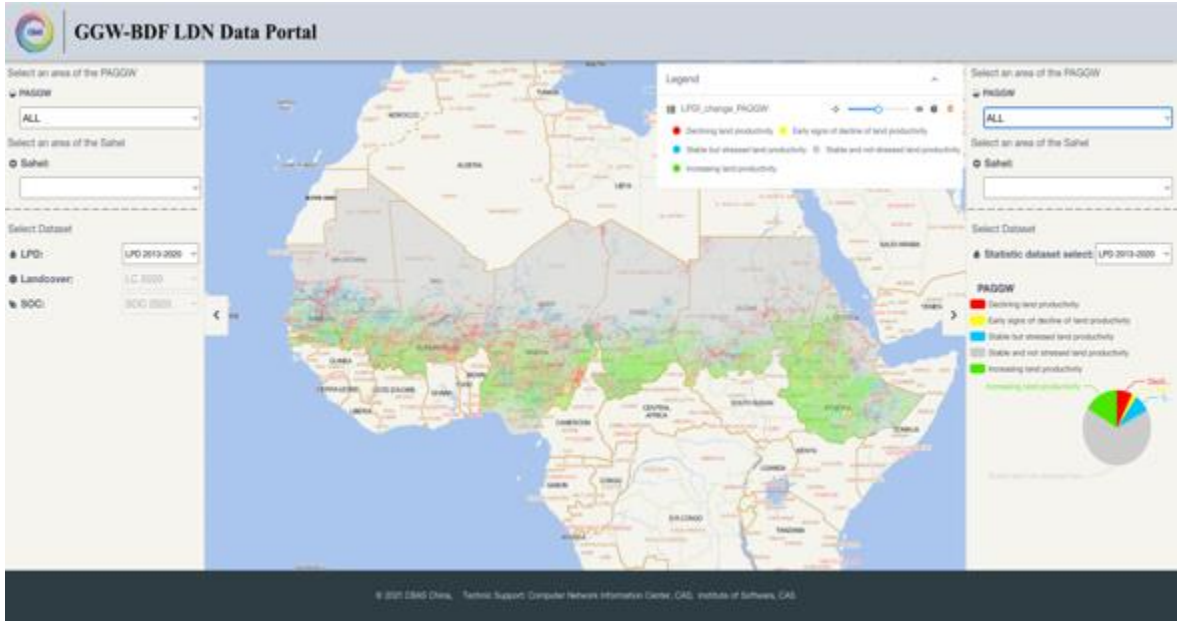
- provide high spatial resolution datasets on land cover, land productivity and soil carbon change to better track the progress of LDN.

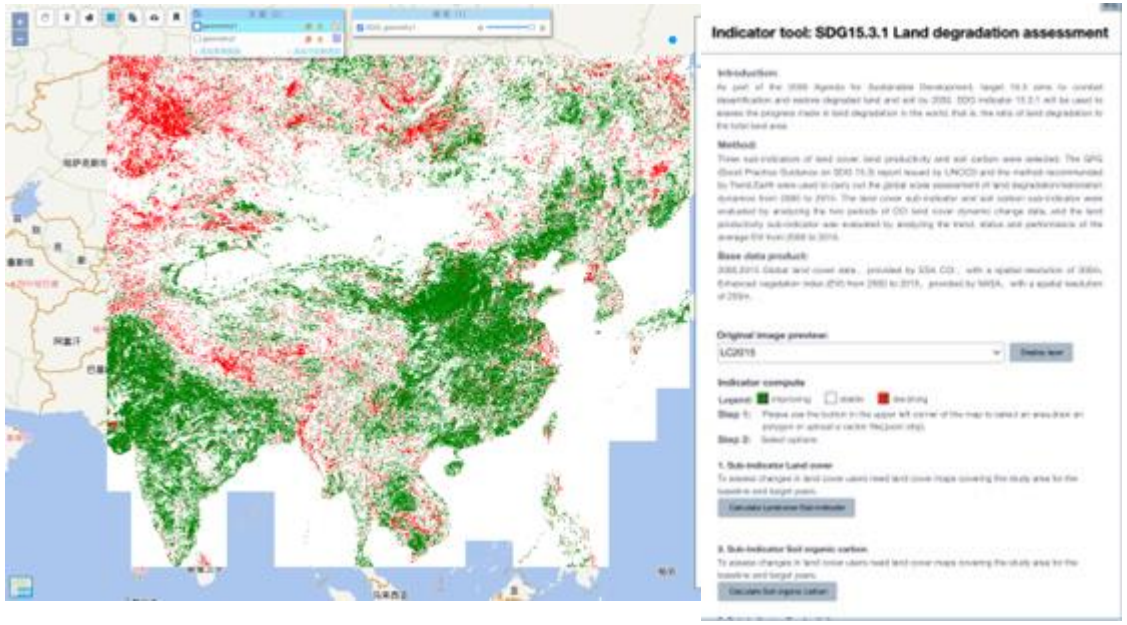




GGW-BDF & SDG15.3.1 LDN

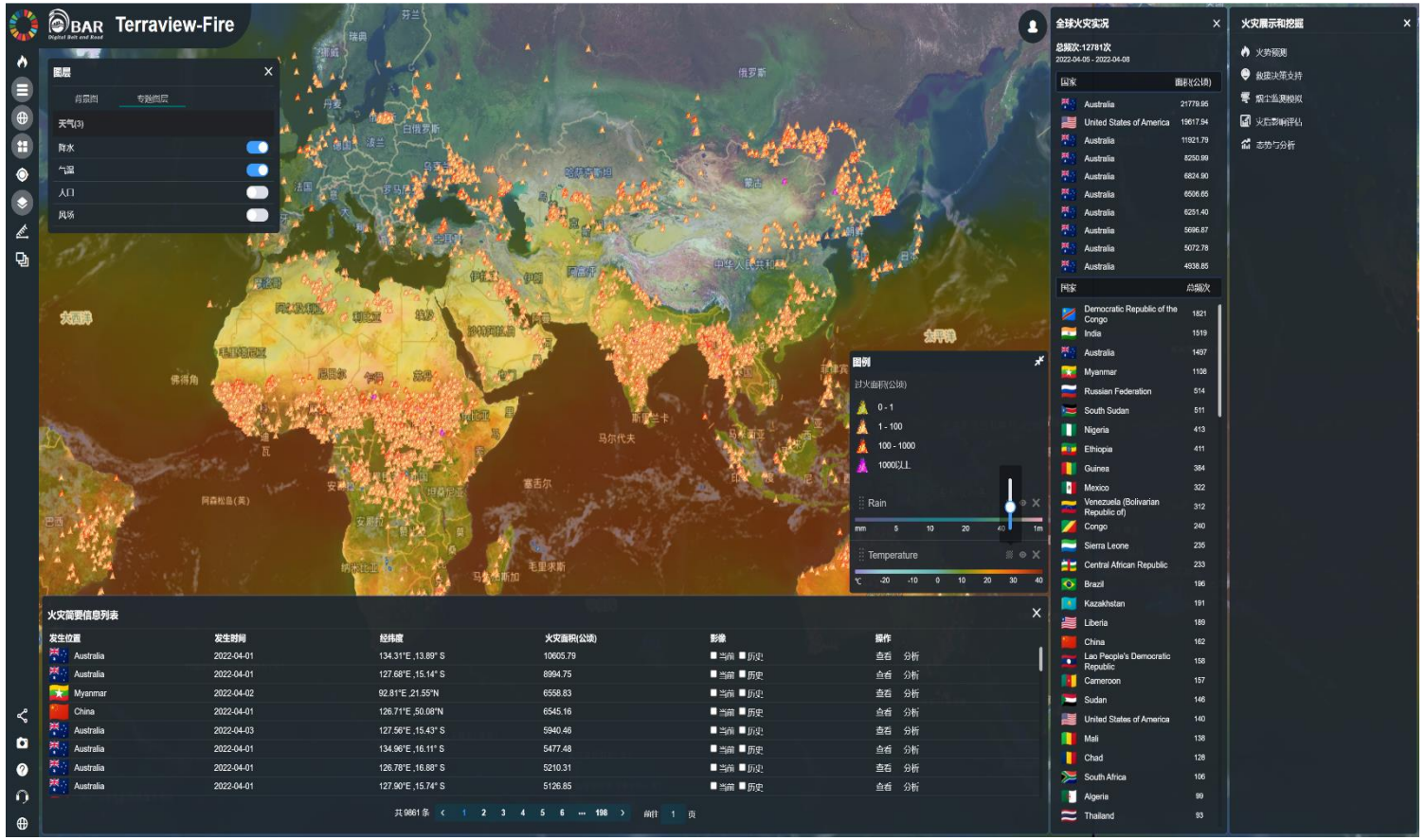
- provide high spatial resolution datasets on land cover, land productivity and soil carbon change to better track the progress of LDN.





Real-time Wildfire Monitoring

- Global Fire Spot Extraction
- Historical Fire Event Analysis
- Burnt Area Mapping
- Weather Data Visualization
- Integrated Fire Spot Analysis
- Real-time Online Computing-supported Fire Monitoring and Decision Support

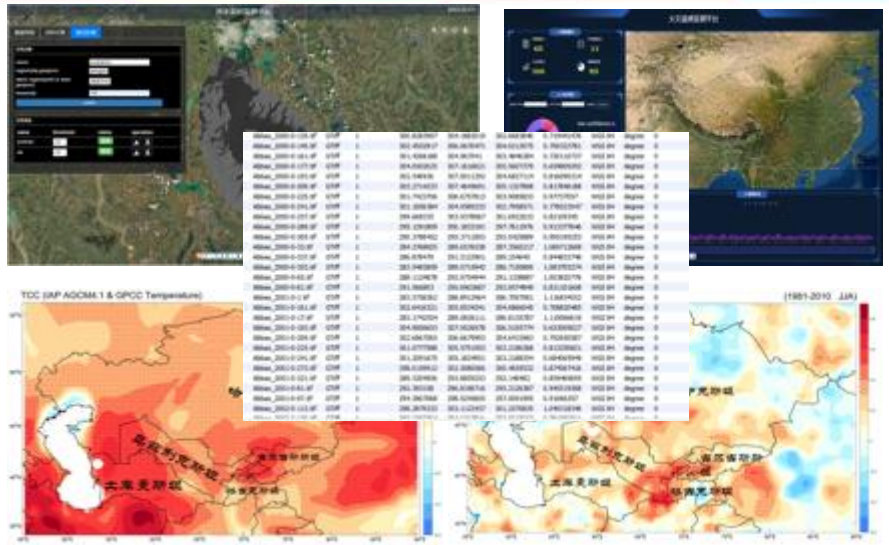


Fire assessment, visualization and decision supported by multi-source data

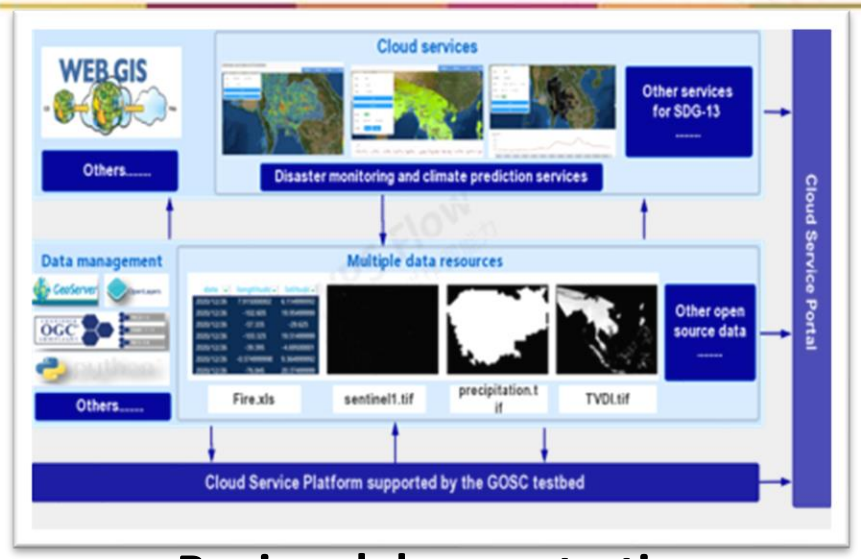
Digital Heritage

Through the DESP and virtual reality technology, 3D reconstruction of the Angkor site was realized as well as 3D virtual tour and interactive functions. The virtual reality system of the Angkor site is built to facilitate the management and conservation of the world heritage site.

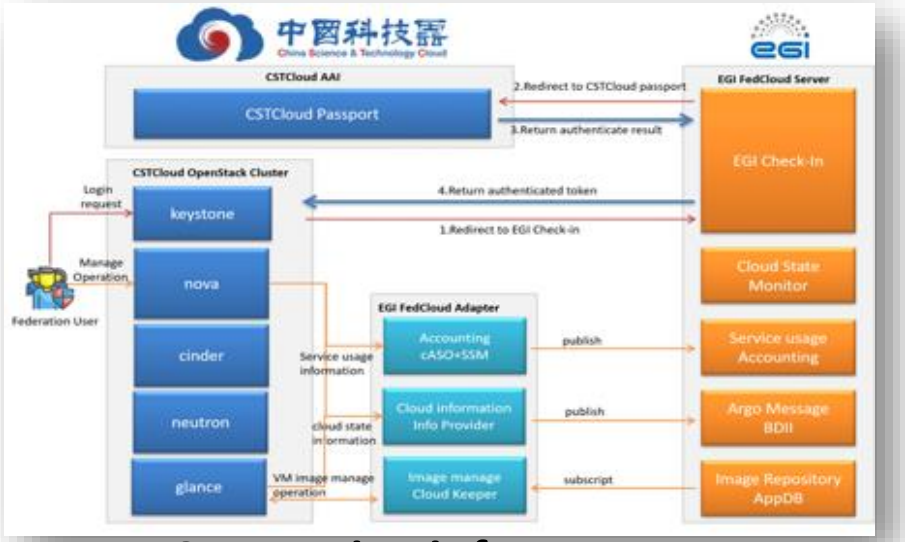
Region climate simulation for SDG13



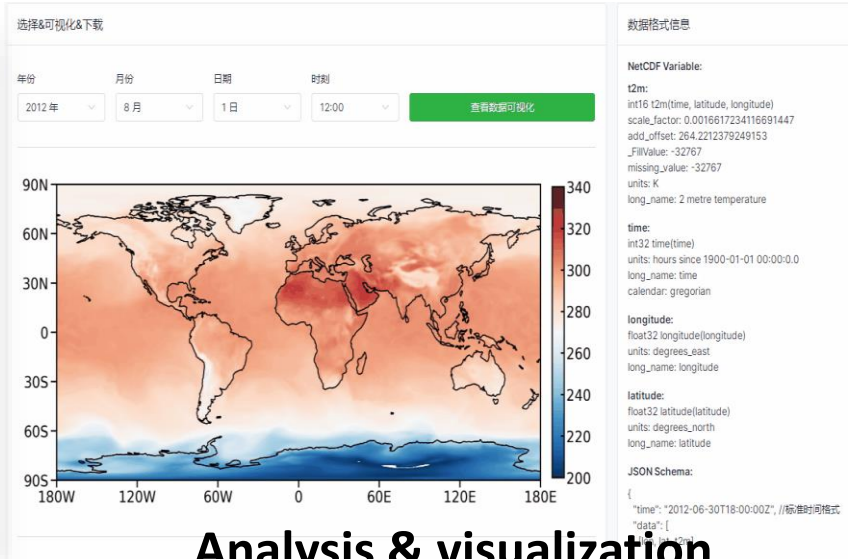
Data resources &models



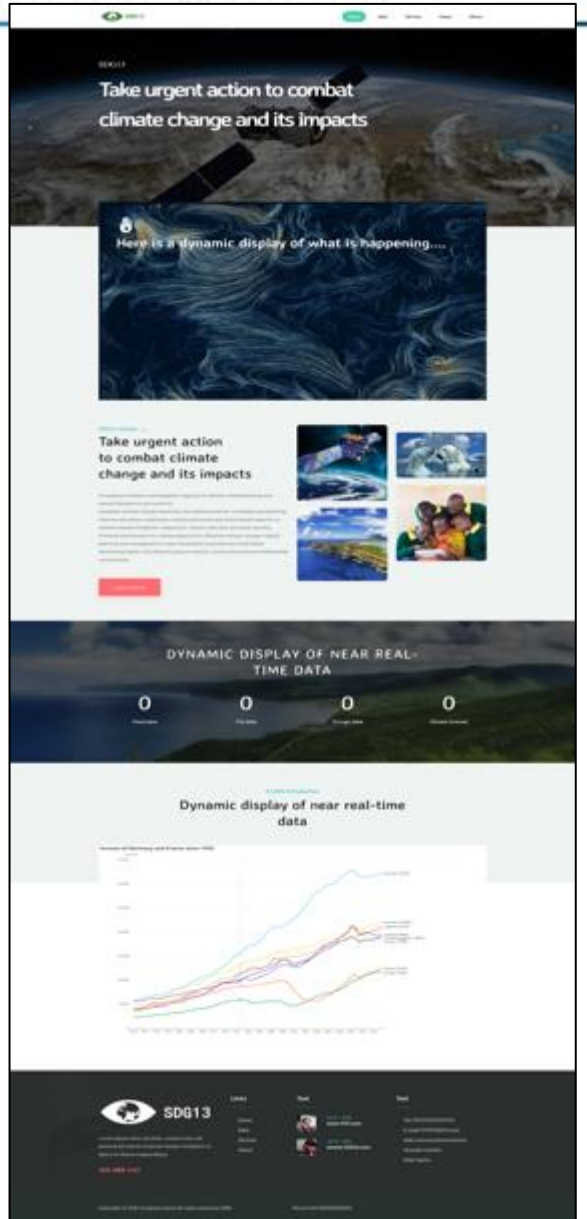
Regional demonstration



Supporting infrastructure



Analysis & visualization



Continental Scale Building Footprint Mapping for Developing Countries

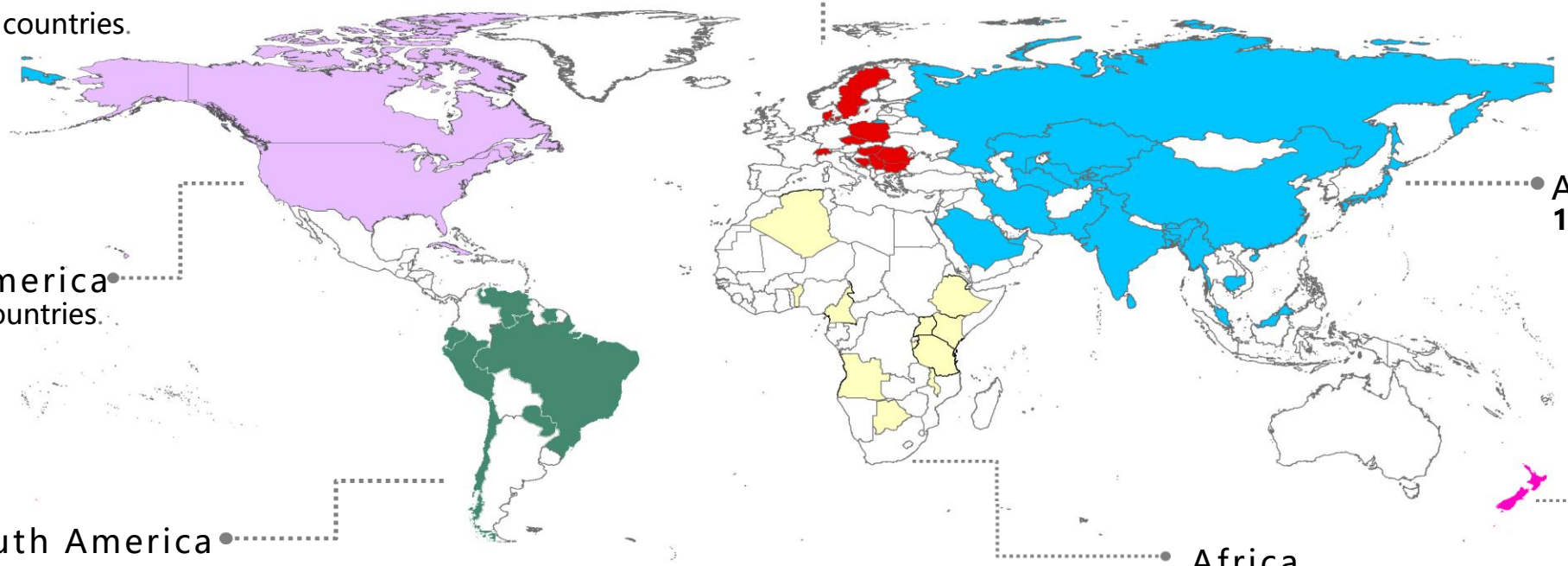
Exsiting Benchmark Data sets for building roof segmentation

Name	Coverage Continent and Country						Coverage City	Resolution (cm)	Coverage Area(km^2)	Num of Building Vector	Data Type
	Asia	Europe	North America	South America	Africa	Oceania					
Massachusetts	-	-	USA	-	-	-	Boston	100	340	254327	RGB
SpaceNet Challenge	China	-	USA	Brazil	Sudan	-	Paris ,Shanghai ,Las Vegas, Atlanta, Khartoum .etc	30-50	5555	685000	8-Band Imagery
Inria	-	Austria	USA	-	-	-	Chicago, Austin, Vienna ,Tyrol .etc	30	810	543672	RGB
WHU	China	Italy Spain	USA Canada	Chile	Egypt	New Zealand	Wuhan ,Los Angles ,Ottwa ,Cairo, Milan, Santiage, Cordoba .etc	7.5-30	1310	22000	RGB
DeepGlobe Challenge	China	France	USA	-	Sudan	-	Las Vegas ,Khartoum, Shanghai ,Paris .etc	31	9623	3020701	8-Band Imagery
AIRS	-	-	-	-	-	New Zealand	Christchurch	7.5	457	220000	RGB

Motivation：

- ❖ Open source image dataset from public maps
- ❖ Large scale building footprint for developing countries
- ❖ Ready for deep learning models

Continental Scale Building Footprint Mapping for Developing Countries



Europe
10 city in 10 countries.

North America
3 city in 3 countries.

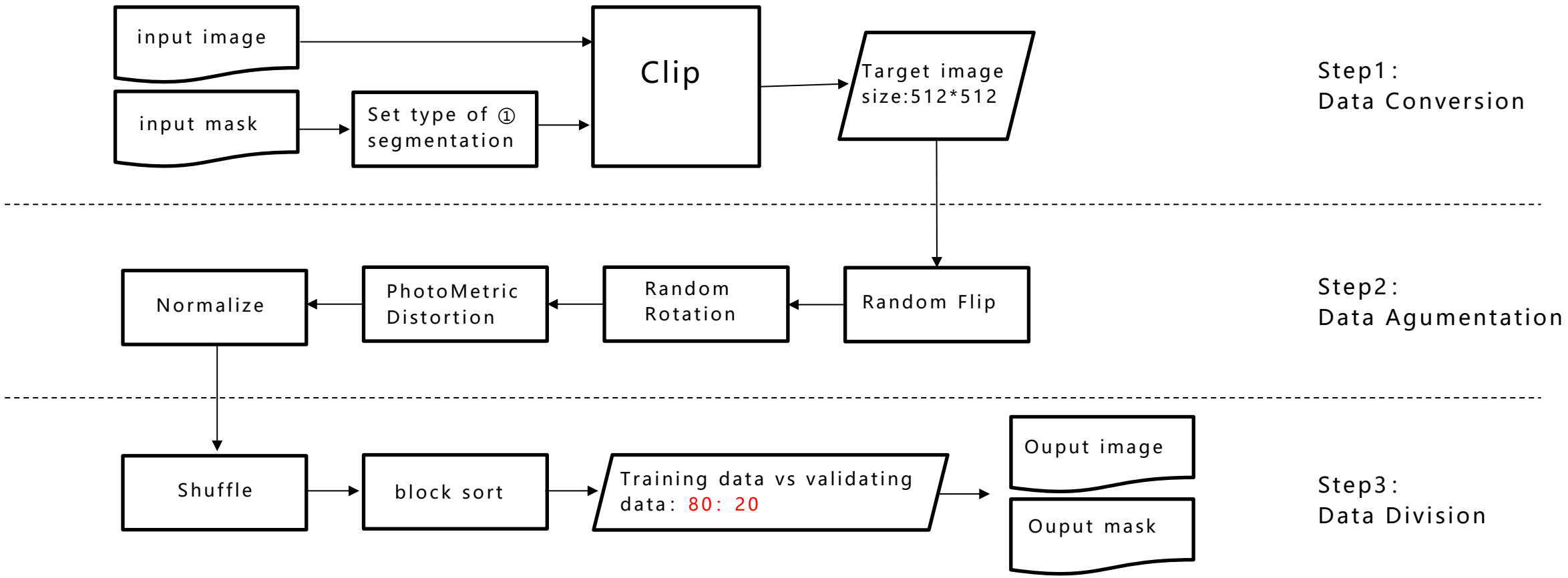
South America
7 city in 7 countries.

Asia
130 city in 20 countries.

Oceania
1 city in 1 countries.

Africa
11 city in 11 countries.

Continental Scale Building Footprint Mapping for Developing Countries

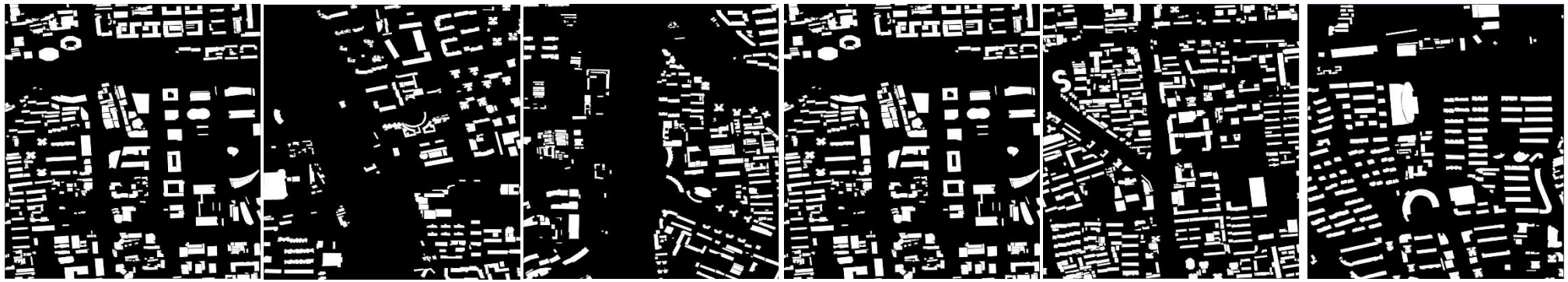


①Set type of segmentation: set the pixel value which represents building to 1, Set the background pixel to 0.

Continental Scale Building Footprint Mapping for Developing Countries

Images

Labels



Baseline Models:

PSP-Net
Pyramid Scene Parsing Network

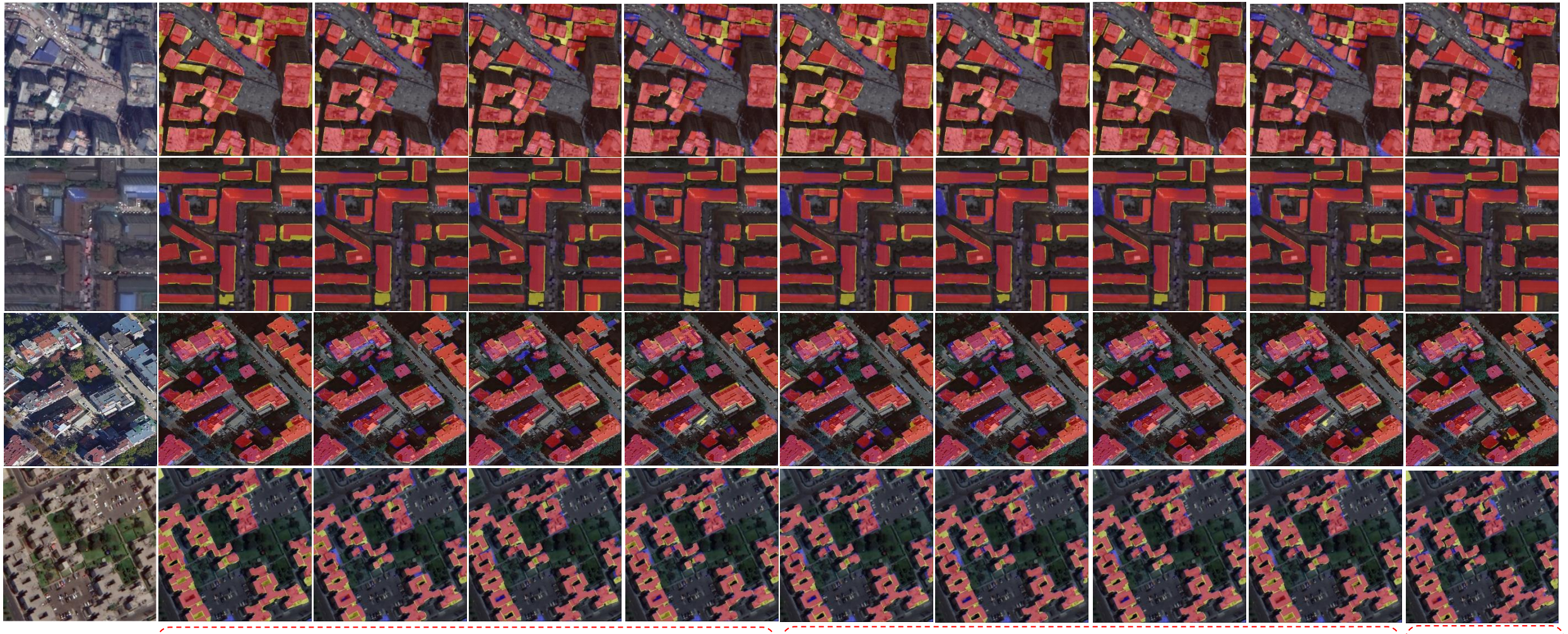
DeepLabV3
Atrous convolution

DeepLabV3+
Effective decoder module

PointRend
Segmentaion as rendering task

SegFormer
Transformer as backbone

Continental Scale Building Footprint Mapping for Developing Countries



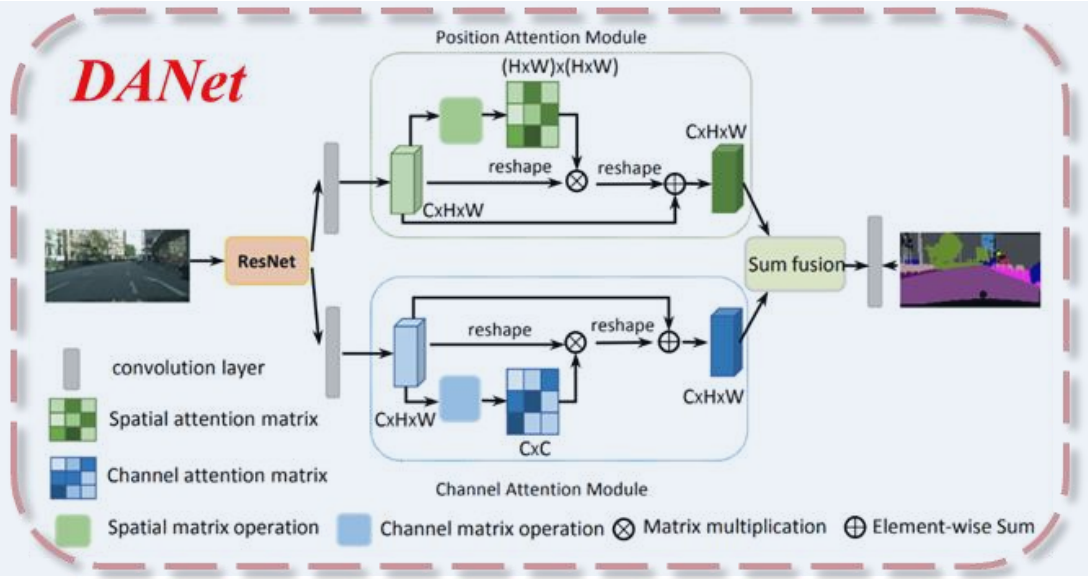
Algorithm assessment at global scale

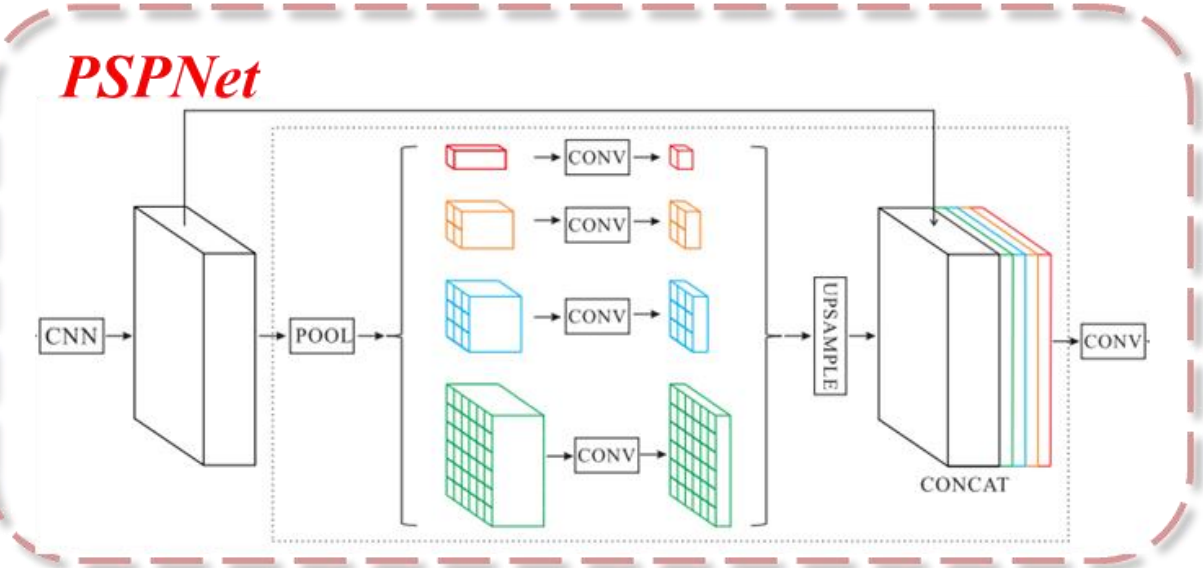
Evaluation of Baseline Models

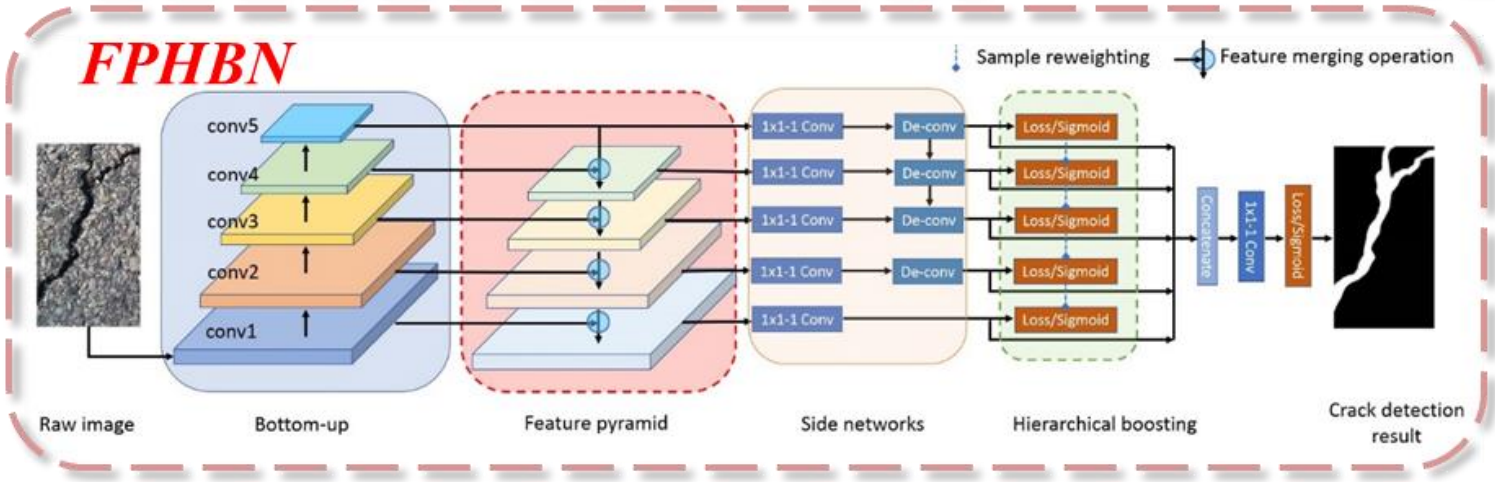
Backbone	Model	Precision	Recall	F-value	PA	mPA	mIoU	OA
ResNet50	PSPNet	90.52	90.64	90.58	96.87	94.80	89.74	96.87
	DeepLabV3	90.6	91.46	91.03	96.68	94.33	89.77	96.68
	DeepLabV3+	91.55	92.32	91.93	97.01	94.9	90.73	97.01
	PointRend	90.91	92.33	91.61	96.9	94.59	90.4	96.9
ResNet101	PSPNet	91.11	92.07	91.59	96.88	94.66	90.36	96.88
	DeepLabV3	90.99	92.5	91.74	96.95	94.65	90.53	96.95
	DeepLabV3+	91.71	92.03	91.87	96.98	94.95	90.66	96.98
	PointRend	91.02	92.26	91.64	96.91	94.64	90.42	96.91
Transformer	SegFormer	90.21	91.56	90.88	96.63	94.16	89.62	96.63

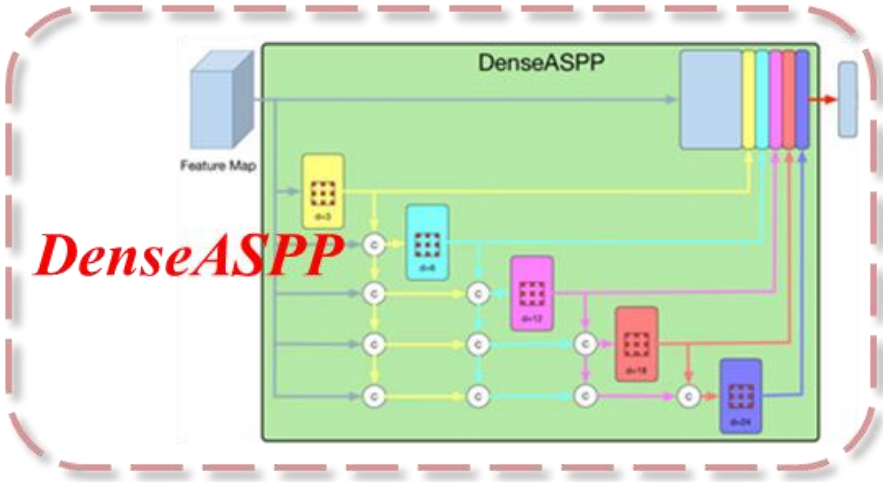
DeepLabV3+ achieved the best results, mIoU> 90%, OA>96% for rest of the models。

Algorithms for crack detection of the dam and rock slopes

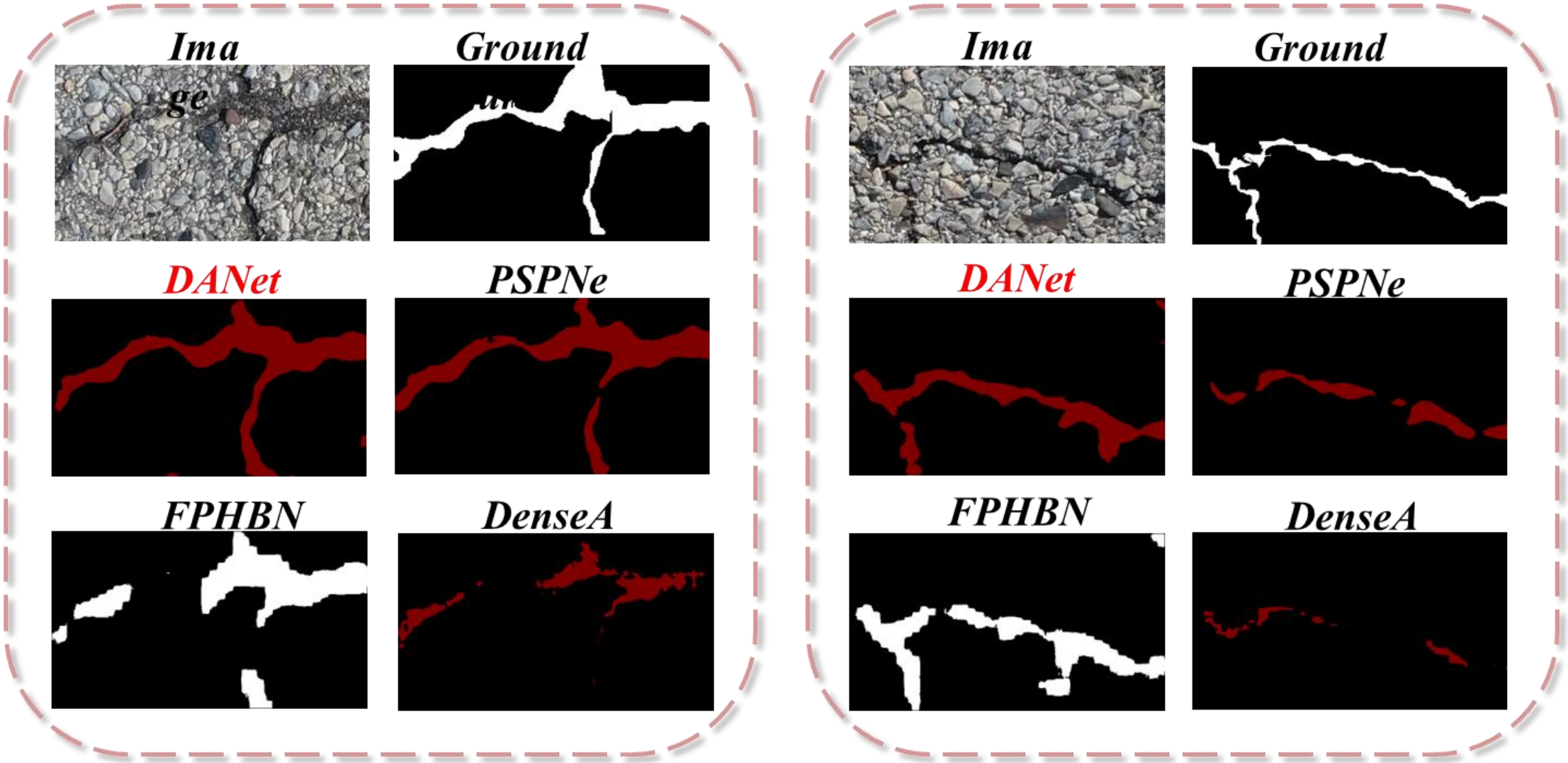




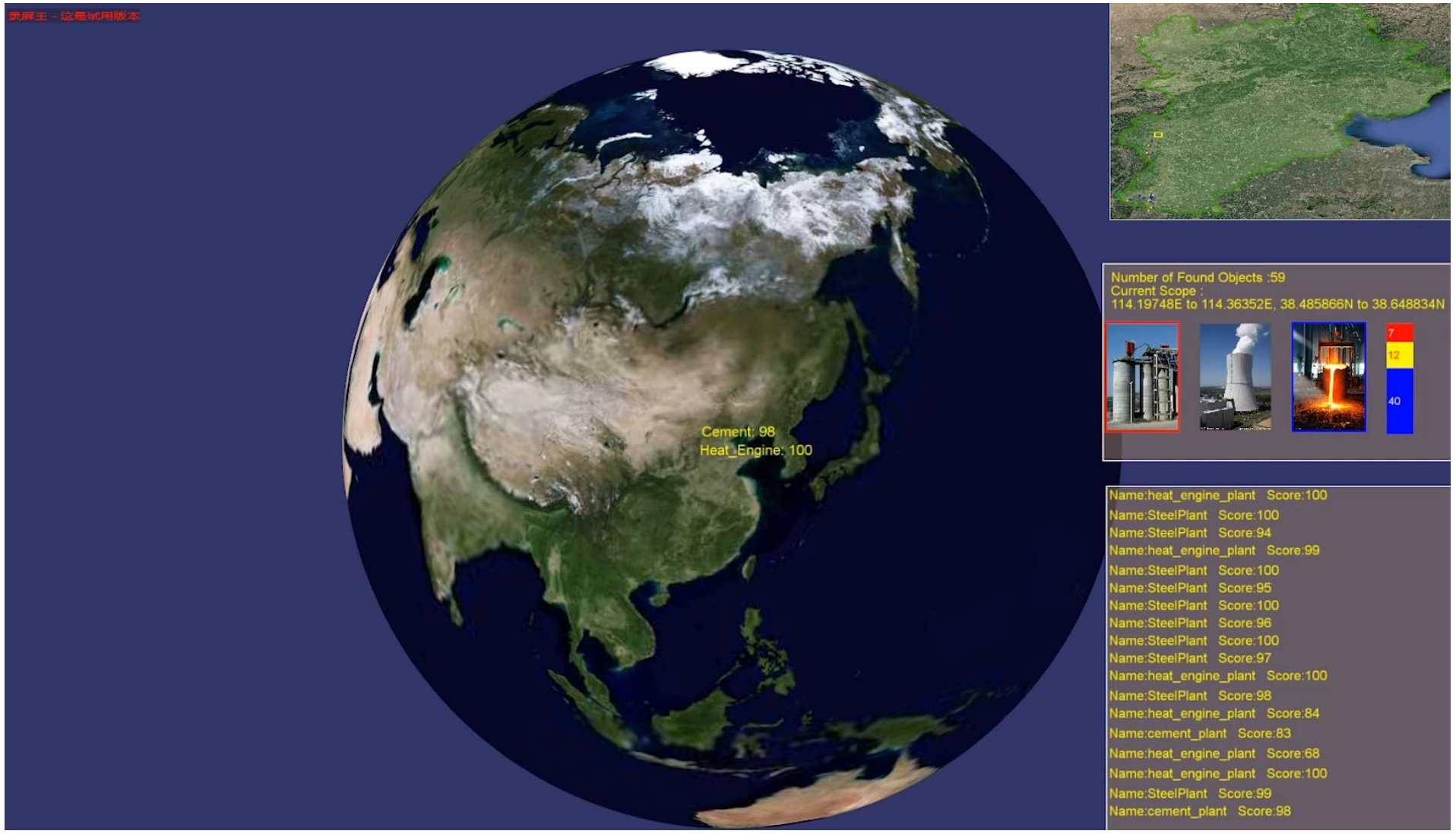




Algorithms for crack detection of the dam and rock slopes



Automatic identification of steel plants / cement plants / power plants



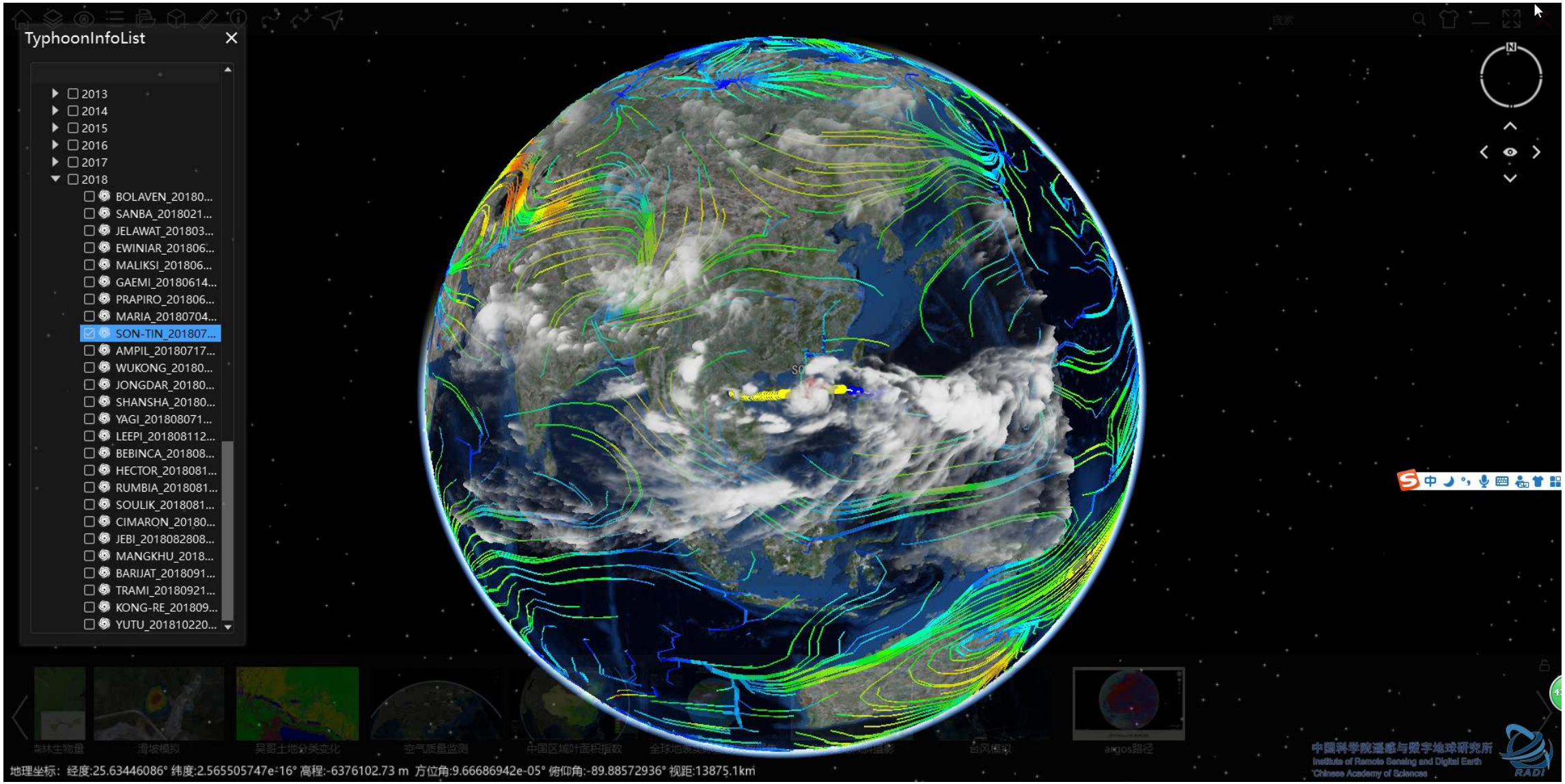
Manual interpretation:
Total: 166
Error: 48
Missing: 21

AI models:
Total: 175
Error: 43
Missing: 7

Coincide: 118
New findings: 21

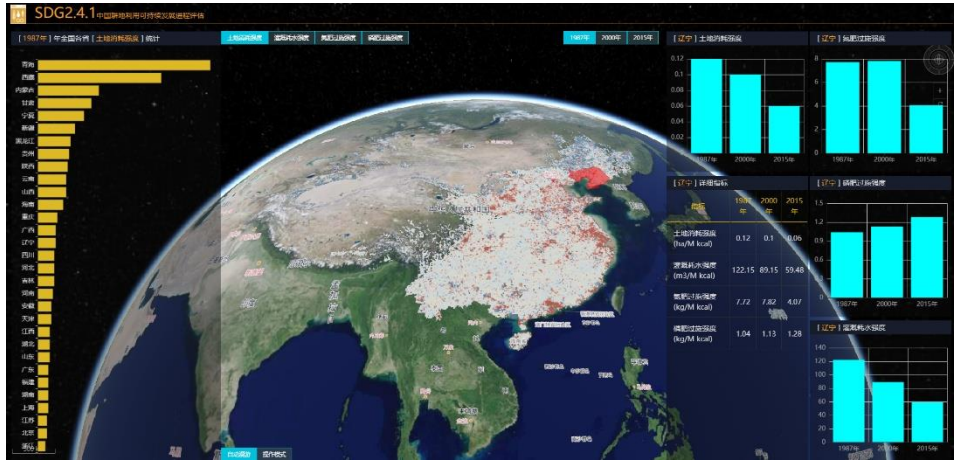
Data: GF1 / 2m
Area: 220,000km²
Time: 2.5h, @4GPU
5-7days (Manual)

Disasters simulation and early warning




SDG11.3.1 全球城镇化进程监测与评估【城市专题】
【曼谷】LCR/PGR指标变化
LCR/PGR
LCR
PGR
1990-1995
1995-2000
2000-2005
2005-2010
2010-2015
【曼谷】人口变化
1990
krung thep (bangkok)
1995
2015
2000
2010
2005
【曼谷】城区面积变化
【曼谷】


DESP EXPLORER


SDG2.4.1 中国耕地利用可持续发展进程评估
【1987年】年全国各省【土地的耗损度】统计
【辽宁】土地的耗损度
【辽宁】详细指标
【辽宁】氮肥过施强度
【辽宁】温室耗水强度

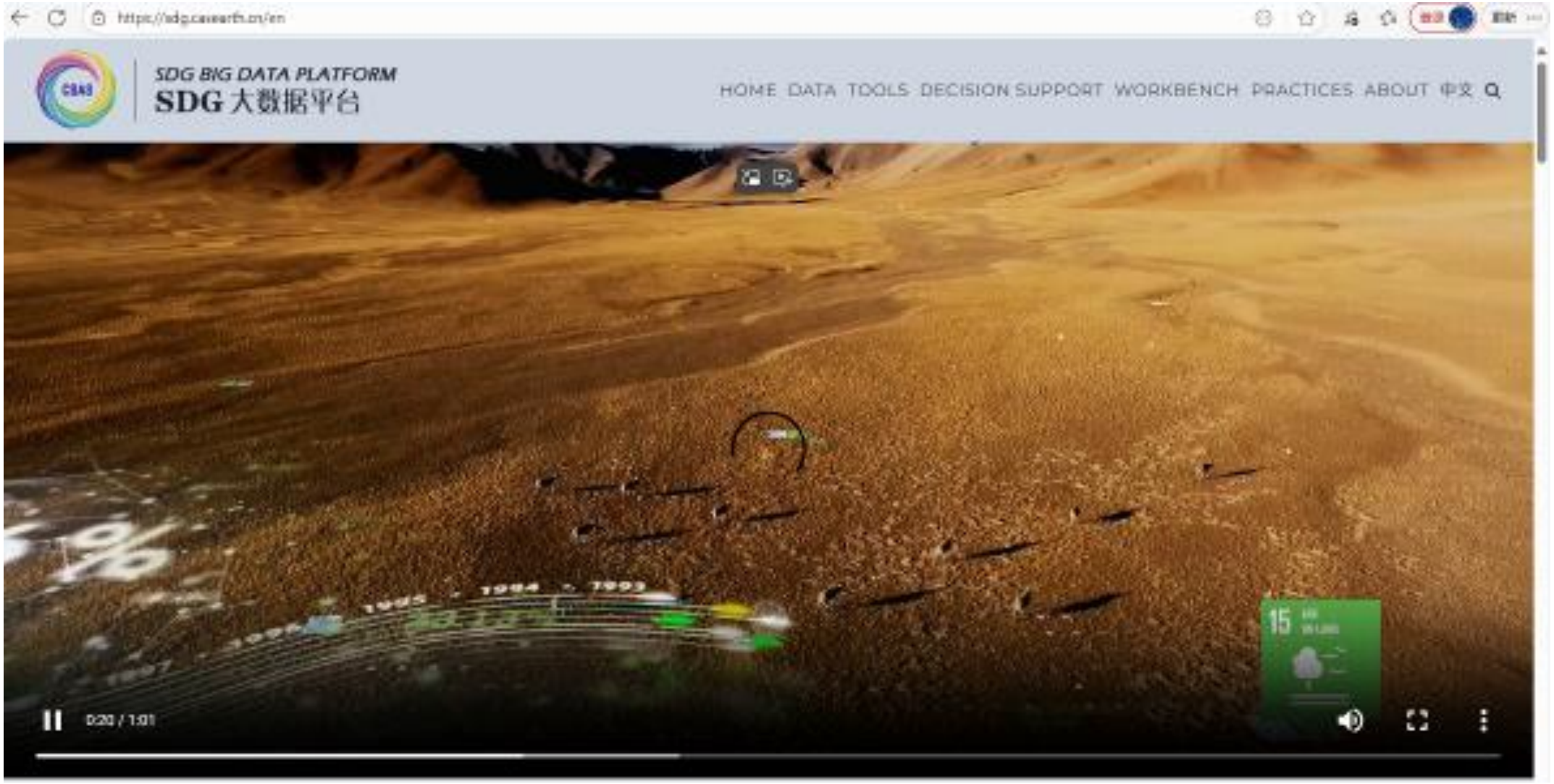
DESP@CBAS is open to all!

CONTENT

- AI-Big Data for SDGs with Cloud Computing
- Infrastructure at CBAS
- Big Earth Data Platform
- Applications for SDGs
- Lab Exercise

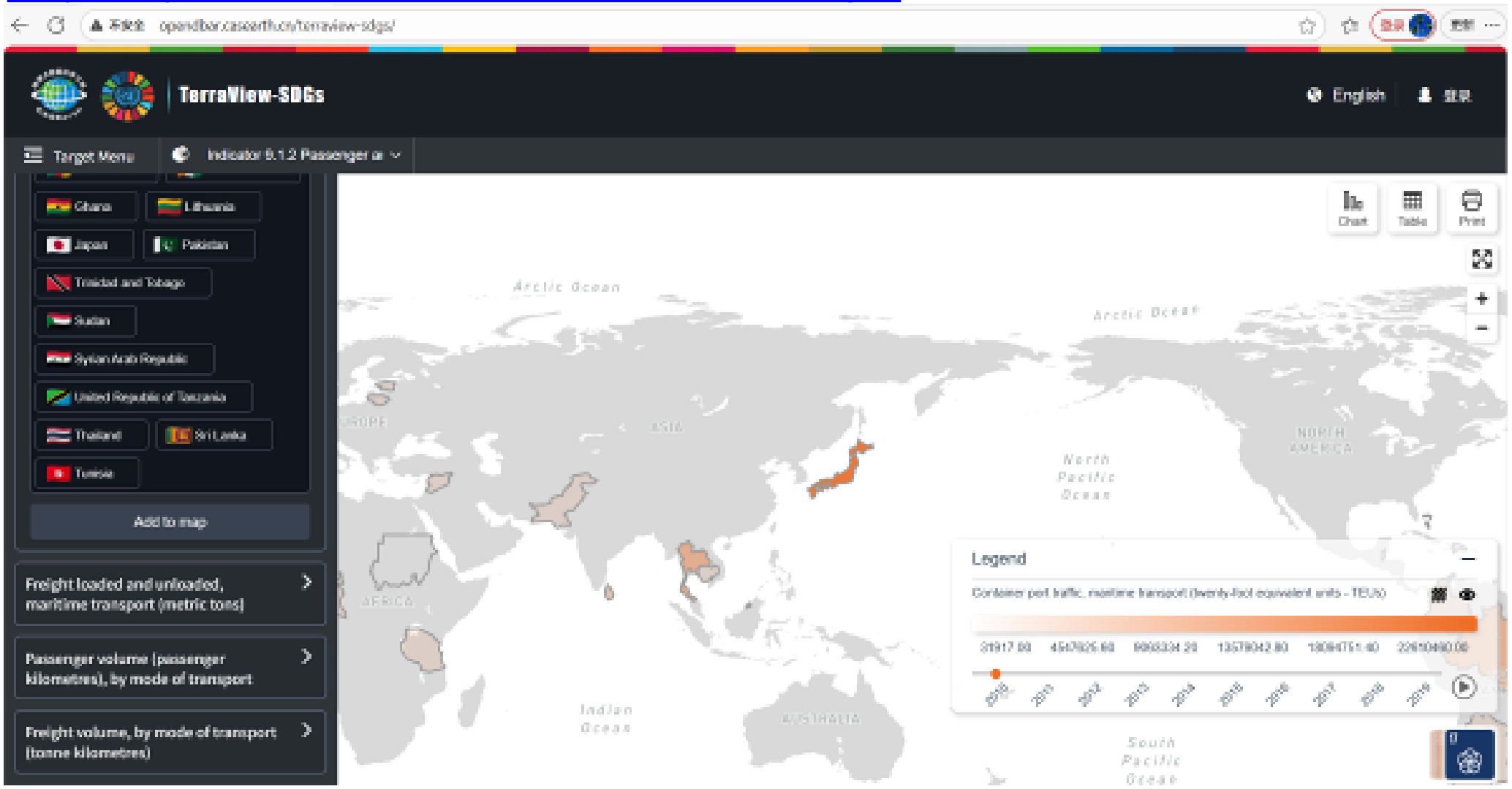
SDG Big Earth Data Platform

https://sdg.casearth.cn/en



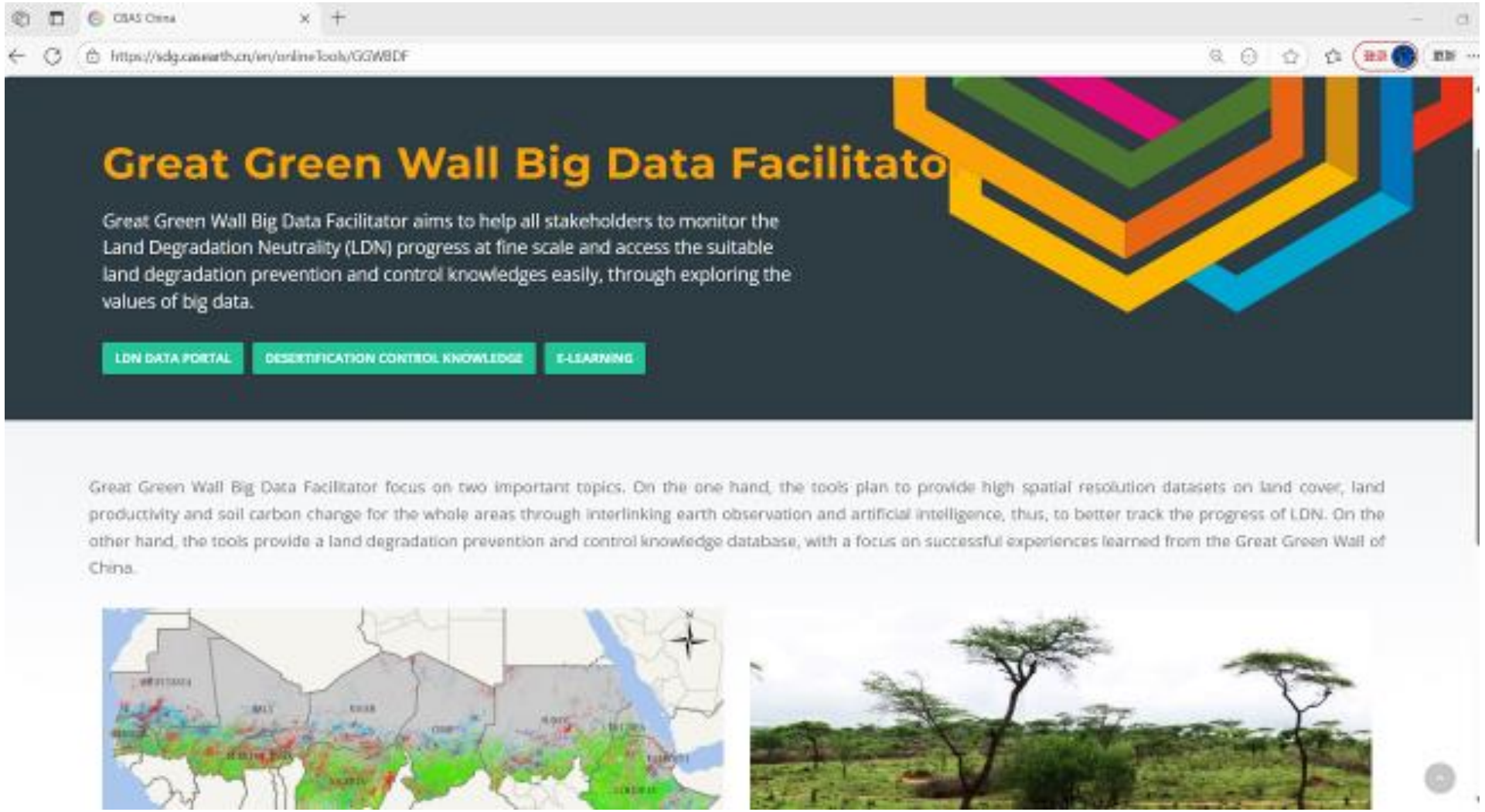
Visualizing global SDGs

http://opendbar.casearth.cn/terraview-sdgs/



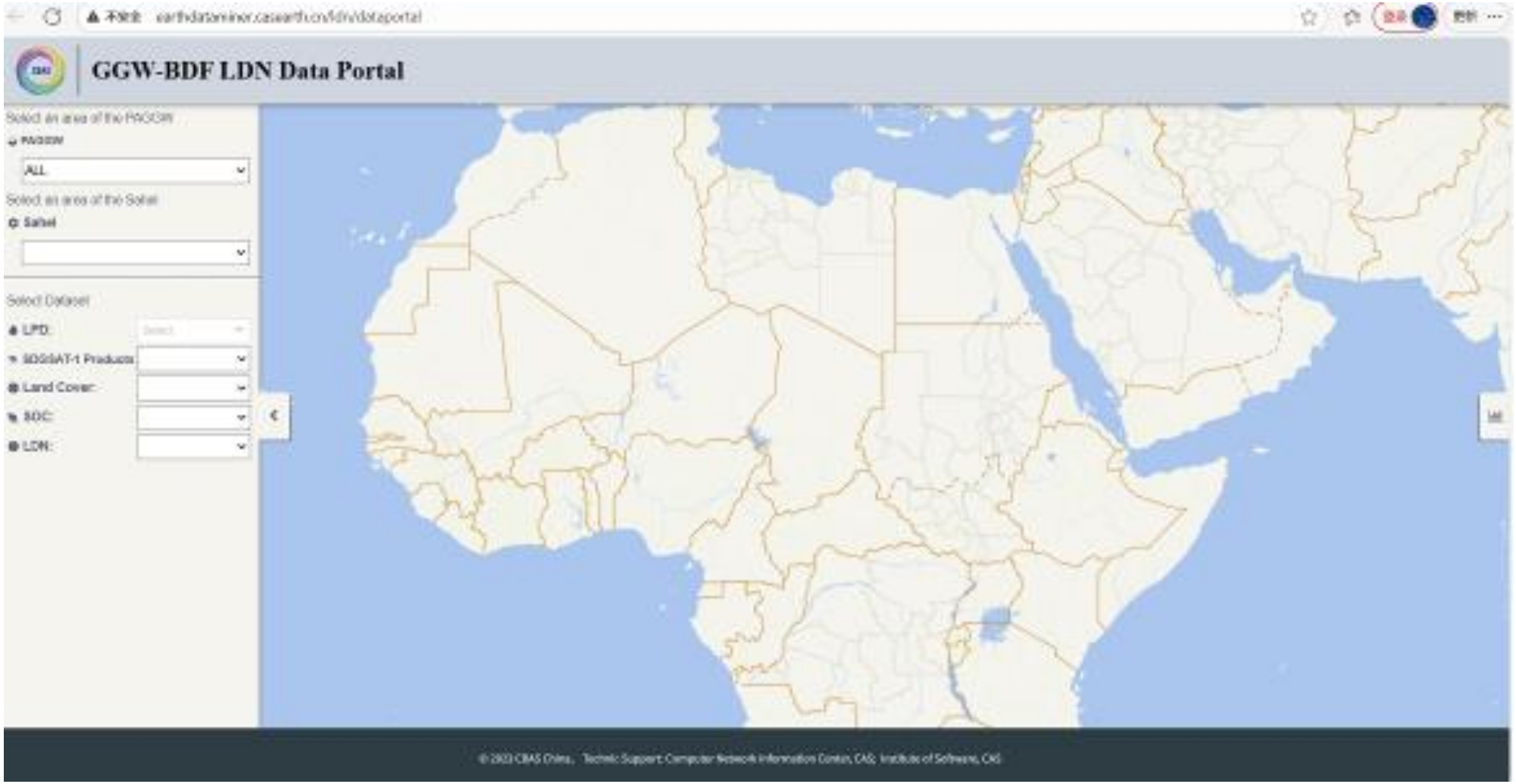
Mapping land degradation with big earth data

https://sdg.casearth.cn/en/onlineTools/GGWBDF



Mapping land degradation with big earth data

http://earthdataminer.casearth.cn/ldn/dataportal

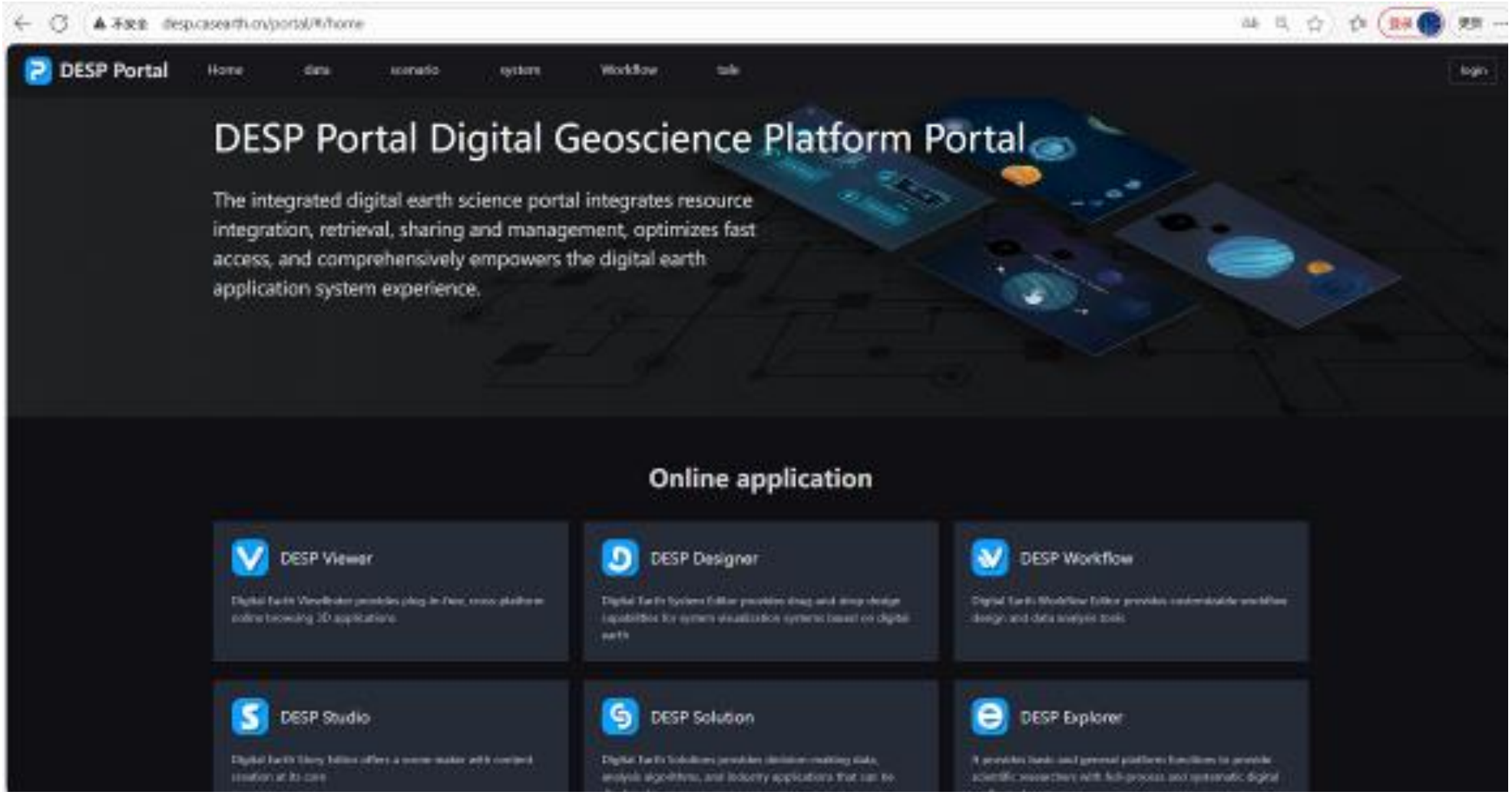


Mapping wildfires with big earth data

http://opendbar.casearth.cn/terraview-fire/

Design your own digital earth apps

http://desp.casearth.cn/portal/#/home



Coding AI model with SDG cloud

Challenge yourself for 1 hour !!!

INTERNATIONAL RESEARCH CENTER OF BIG DATA
FOR SUSTAINABLE DEVELOPMENT GOALS
可持续发展大数据国际研究中心

Thanks

No.9 Dengzhuang South Road, Haidian District, Beijing 100094, China
Tel: +86-10-82178985
Fax: +86-10-82178980
Http//:www.cbas.ac.cn