DIAS Platform Contributing to Open Science in Earth Environmental Informatics

Masaki Yasukawa[1]*

[1]* *Earth Observation Data Integration and Fusion Research Initiative, the University of Tokyo,*
4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505, Japan
Email: yasukawa@iis.u-tokyo.ac.jp

Summary. Data Integration and Analysis System (DIAS) is one of data centric cloud. DIAS collects and integrates massive and diverse data such as earth observation data and socio-economic data, and various applications on DIAS platform provide beneficial information to solve societal issues such as environmental problems in Japan and overseas. This article gives the concept of DIAS at first. As a case study on DIAS's contribution to open science, I introduce a data sharing system for flood forecasting in Sri Lanka, and data collecting systems for citizen science of butterfly and oriental stork.

Keywords. DIAS, open science, data centric cloud, data sharing, citizen science.

1. Introduction

Data Integration and Analysis System (DIAS) started from 2006. The goals of DIAS are to collect and store earth observation data; to analyze such data in combination with socio-economic data, and convert data into information useful for crisis management with respect to global-scale environmental disasters, and other threats; and to make this information available within Japan and overseas. The prototype of DIAS was developed in 2010. Then, phase II of DIAS from 2011 made further advancement and expansion to apply DIAS as a social and public infrastructure. The current project of phase III has started since 2016 with the aim of its practical operation.

2. DIAS

The three systems that comprise DIAS are shown.

- Infrastructure system: integrates Earth and local observation data and numerical models, socio-economic data, and other massive datasets relating to the global environment.

- Application development: implements storage, search, analysis, visualization, and other data-related functions on the infrastructure system to provide scientific knowledge and resolve global environmental and societal issues.

- Research and development (R&D) community: DIAS provides a close R&D environment that allows domain scientists who study the environment and IT experts to promote joint research and development through cooperative work, planning, and production.

DIAS project is unique, because a community to support application development on data infrastructure is established in addition to the construction of data infrastructure.

3. DIAS Platform

DIAS provides the platform to process the global environment data. When the domain researchers install their programs in DIAS, and they use real-time data archived on DIAS, the application of real-time processing is easily realized. The applications related to an open science on the DIAS platform are introduced as follows.

3.1 Data Sharing System for Flood Forecasting in Sri Lanka

In Sri Lanka, a large flood occurred in late May 2017, and many residents were sacrificed. For the secondary disaster precaution and the basin reconstruction, a web-based data sharing system

for flood forecasting has been developed in June 2017, and the data providing has been started to the local stakeholders.

This system has data processing tools and visualization tools. The data processing tools include data collection, data correction, and real-time prediction. The visualization tools can display various real-time data such as rain gauge data, satellite data, rain map, rain forecast and flood prediction.

The local stakeholders use this system for flood monitoring. Also, this system is useful for teaching material for technology transfer. For more effective flood prediction in future, they will consider the suitable flood prediction model, and construct the new system themselves.

3.2 Data Collecting System for Citizen Science of Butterfly

The butterfly is familiar with citizen, and suitable as the index of global warming and urbanization.

A data collecting system for citizen science of butterfly in Tokyo has been constructed in 2009. The participant uses the data upload tool, and uploads the monitoring data with photograph. Data manager of butterfly expert uses the data quality control tool, and do data cleansing. After the cleansing, the data is open to the public by data visualization tool.

In 9 years, about 50,000 records were uploaded. A lot of findings such as the distribution expansion of southern butterfly, the distribution expansion of alien species, and existence of the species on the red-list were brought by the power of the citizens.

3.3 Data Collecting System for Citizen Science of Oriental Stork

The last one of wild oriental stork in Japan was died in 1971, due to farmland consolidation, disappearance of wetlands by river refurbishment, and the use of pesticides. By the activity of feral population reproduction with breeding and releasing, about 100 oriental storks inhabit satoyama landscape in Japan now.

For contributing to the monitoring activity of oriental storks, a data collecting system for citizen science of the stork has been developed by enhancing the collecting system of butterfly in April 2018. In this activity, the feature is to identify the individual number of the stork. About 1,800 records with photograph were collected from April to August 2018. The habitat map and the safety confirmation list were made from the dataset. Using the location, the behaviour and the food, the clarification of the history of life and personality will be expected in future.

4. Conclusions

DIAS is not only a data repository, but also an application platform. The field of the application on DIAS platform is wide range. Also, some of applications are contributing to open science for earth environment. Especially, data collecting systems for citizen science are the practical examples of open science. By collecting more data of butterfly and oriental stork and developing the visualization tools, a lot of new findings on ecology will be expected in future.

References

1. Kawasaki, A., Yamamoto, A., Koudelova, P., Acierto, R. A., Nemoto, T., Kitsuregawa, M., Koike, T., Data Integration and Analysis System (DIAS) Contributing to Climate Change Analysis and Disaster Risk Reduction. *Data Science Journal*, 16, 41, 1-17, 2017
2. Yasukawa, M., Ikoma, E., Nemoto, T., Rasmy, M., Tsuda, M., Ushiyama, T., Tamakawa, K., Koike, T., Kitsuregawa, M., Prototyping a Data Sharing System for Flood Forecasting: A Case Study on Sri Lanka, *3rd International Symposium on Big Data Analytics in Science and Engineering (BASE 2017)*, 2017
3. Ikimoni, http://butterfly.diasjp.net/ [accessed on: September 2018] (in Japanese)
4. Citizen Science of Oriental Stork, https://stork.diasjp.net/ [accessed on: September 2018] (in Japanese)