Development of Data Integration and Information Fusion Infrastructure for Earth Observation

Mikio Takagi
Shibaura Institute of Technology
3-9-14 Shibaura, Minato-ku, Tokyo 108-8548, Japan
Email: mikio@sic.shibaura-it.ac.jp

Masaru Kitsuregawa
University of Tokyo
4-6-1 komaba, Meguro-ku, Tokyo 153-8505, Japan
kitsure@tkl.iis.u-tokyo.ac.jp

Ryousuke Shibasaki
University of Tokyo
4-6-1 komaba, Meguro-ku, Tokyo 153-8505, Japan
shiba@paddy.iis.u-tokyo.ac.jp

Seishi Ninomiya
National Agricultural Research Center
3-1-1 Kannondai, Tsukuba, Ibaraki 305-8666, Japan
snino@affrc.go.jp

Toshio Koike
University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
tkoike@hydra.t.u-tokyo.ac.jp

Abstract: The 10 Year Implementation Plan for a Global Earth Observation System of Systems (GEOSS), which was endorsed at the Third Earth Observation Summit in Brussels in February, 2005, emphasizes the importance of data management facilities for diverse and large-volume Earth Observation data from inhomogeneous information sources. A three year research plan for addressing this key target of GEOSS has just approved as the first step by the Japanese government.
The goals of this research are,

(1) to develop a data management core system consisting of data integration and information fusion functions and interoperability and information service functions;
(2) to establish data and information flows between data providers and users;
(3) to promote application studies of data integration and information fusion, especially in the fields of weather forecasting, flood forecasting, agricultural management, and climate variability and changes.

The research group involves leading scientists on information science and technology, who have been developing giant data archive servers, storage area networks, metadata models, ontology for the earth observations. They are closely cooperating with scientists on earth sciences, water resources management, and agriculture, and establishing an effective collaborative research framework.

Keywords: Global earth observation, Data integration, Information fusion, Large-volume data, Interoperability

1. Introduction

A better understanding of the global earth system and an advanced forecasting and reaction capability for the global changes, using earth observation data, are strongly expected in today's international community. Earth observation involves natural and human related earth subsystems and further interaction and variations. It also needs to handle extremely diverse data obtained from various information sources. More chances to utilize satellite observation data and the sudden increase in a numerical model output are taken into consideration, the data dealt with by earth observation

serves as super-large scale inevitably, and it estimates that its data amount will reach a several tens of peta (peta is ten to the 15th power) byte by the year of 2010.

Thus, in order to draw out scientifically useful knowledge from a large-scale heterogeneous earth observation data and to effectively utilize the acquired information for such as making policies, and decision making in crisis and resource managements, the system which makes easy to access the data , easy to analyze the data, and easy to use the data, needs to be built. Thus, through the systematized data collection, the strategic control of the data with considering the life cycle of data, the effective data integration, and the acquired information fusion, those data will be converted into the scientifically and socially useful data, and further more the system construction for the international data sharing is required. Moreover, as regard to the relatively small –sized data, it is important to build a distributed data system connected in the network, in consideration of future data integration and centralized data control, trying to unify the data format and the data protocol.

2. Purpose

In this research, the mutually cooperative organizations and groups from several different fields, such as the information-science-and-technology, the earth observation, disaster and agriculture, will work together to try to integrate the earth observation data effectively and efficiently, and to fuse the information, then to aim to develop a demonstrative prototype of the data system to be shared internationally.

3. Contents of Research

In this research, we develop the following two core systems; (1) Development of large scale earth environmental data integration and information fusion system and (2) Development of Mutual Use and Information Service System, then for the input data into the systems above, we also aim for (3) Research Development on Earth Observation Data Collection and Data Quality Control, and (4) Research and development on advanced information application technology which will make the integrated and fused earth observation information to be applied to the public interest field.

The outline of each sub-theme research is described as indicated below.

3.1 Development of large scale earth environmental data integration and information fusion system

The integrated archive of earth environment data and construction of a flexible information fusion mechanism play a very important role as a research base in the earth environment field. Earth environment research in the 21st century will be considered that the tool which looks for the integrated warehouse of huge and various data free becomes the time of deciding the success or failure of original research. In this sub-theme, it is going to build timely the firm data base which can respond to various research needs. These data base research and development overly consist of development of the core technology of huge archive and data assimilation, advanced visualization mining, XML schema, high order network, and sensor fusion.

This sub-theme is consisting of next five sub-sub-themes.

- Development of large scale data archive/storage system and data integration infrastructure.
- Development of Data Assimilation System.
- Development of Advanced Data Visualization and Mining System.
- Development of advanced network infrastructure for huge scale data collection and distribution.
- Development of data storage and fusion framework for active data sensor network.

3.2 Development of Mutual Use and Information Service System

In many fields, such as agriculture, disaster, ecosystem, and biodiversity, communalization of a technical term, or a classification system, standardization of the description method of a place, etc. have not been in much progress, let alone structuring of data. And this delay has been an obstacle of mutual use of these data, or distributed use. Moreover, development of the distributed use system which can have a more flexible and easier response to various use needs, are also desired about satellite observation data. So, in this research, we try to develop the ontology registry which can record, manage, and compare ontology related information, such as a data dictionary, a classification system, and a thesaurus. and various information share / use support services, such as high order data processing, analysis support, etc. of the metadata design support using it, structure support of data contents, text mining support, the data analysis between space-time, etc., are built. And agricultural field data and satellite observation data are taken up as a case study to examine this applicability.

This sub-theme is consisting of next four sub-sub-themes.

- Development of Ontology Registry (Recording, Comparing, Editing System of Ontology Information).
- Information Sharing and Access Support Service using Ontology Registry.
- Distributed data fusion system in agriculture.
- Development of Distributed Satellite Data Integration System.

3.3 Research Development on Earth Observation Data Collection and Data Quality Control

The datasets provided by the earth observation are satellite observation data and its high order processing products and ground (land, ocean) observation data. The satellite observation data we emphasize the long term observed data collected by the JAXA and the international cooperation between satellite organizations.
Regarding the filed observation data we archive the flood observation data, marine research data and research ground observation data collecting by the individual organizations. By putting these data in a database, we make a environment of research and development to use these data integrally.

This sub-theme is consisting of next five sub-sub-themes.

- Improvement of Satellite Data Collection and Data Processing Methods.
  - Development of Multidimensional Data Integration for Satellite Data.
  - Development of satellite data reception, management and archiving system.
- Satellite Data Set processing.
- Integration of hydrometeorological data.
- Integration of Oceanographic Data.
- Integration of research in-situ data.

3.4 Research and development on advanced information application technology

For a more comprehensive understanding of the global earth systems, it is indispensable to give the capability of analyzing the complexity system by consolidating the observation data and social economic data, and also by introduction of new techniques such as data assimilation and data mining. In this context, the observation data means the data regarding the response in the human area toward about the physical system of atmosphere, land, and ocean, the living thing system, and the chemistry system. The prediction with more high- accuracy will be attained by the comprehensive understanding above.

Given to meet the demand of contributions from the earth observation field in decision-making and policy-making in the public profits field (such as disaster, health, energy, climate, water resources, weather prediction, agriculture, an ecosystem, and biodiversity) , it is required that the technology for translating the earth observation data into some useful information in accordance with a real management needs and the technology for offering the information about the local future trend depending on the local characteristic are especially developed for Japan and Asian regions.

This sub-theme is consisting of next four sub-sub-themes.

- The earth environmental data integration and information fusion for improvement of numerical weather forecasting.
- Application Research of Earth Observation Integrated Data for Water Management and Its Fused Information.
  - Development of Utilization System in Water Management Application.
  - Development of atmosphere, land and river channel coupled model.
- Utilization of integrated earth observation data and information fusion for optimal agricultural management.
  - Totalized data assimilation and integration methods for crop managements.
  - Soil-Vegetation Model and Its Model Integration.
- Utilization of Earth Observation Integrated Data and Its Fused Information to Advance Climate Information.
  - Study on the verification and application of JMA Climate Data Assimilation System.
  - Analysis of earth's global warming phenomenon with use of the integrated database.

4. Implementation Structure

The following two groups will work and collaborate together and build an organic cooperative and comprehensive research organization. One group is the cutting-edge research group in the field of information science and technology, such as super-huge archive and ontology, who has been long working on an integrative use of earth observation data and information fusion for years, in order to use inhomogeneous large-volume earth observation data efficiently and effectively.

Another group is the research consortium for data provider and data use about the earth observation, the weather and climate prediction, the nation's land and agricultural management, etc.

The implementation structure is shown in the Table 1.

5. Advantage Which Research Progress and Result Bring About

The 10 Year Implementation Plan for a Global Earth Observation System of Systems (GEOSS) is proposed by the initiative of Prime Minister Koizumi in the Group of Eight Summit held in Evian in 2003 and Japan as one of the Co-chairs has consistently led the intergovernmental working group. In correspondence with this proposal, "The promotion strategy of earth observations" has been selected by the Council for Science and Technology Policy. We believe the result achieves a great contribution in each field of the correspondence to the weather, agriculture, water management, climate changes, and disasters not only in Japan but in Asian nations.

ACKNOWLEDGEMENTS

This research will be implemented with the support of the Ministry of Education, Culture, Sports, and Science and their Science and Technology Promotion Research Fund. We really appreciate their assistance.

Table 1. Implementation structure for research

Research Theme	Institute Name	Researcher Name
1. Development of large scale earth environmental data integration and information fusion system		
1.1 Development of large scale data archive/ storage system and data integration infrastructure	Institute of Industrial Science, The University of Tokyo	Masaru Kitsuregawa
1.2 Development of Data Assimilation System	School of Engineering, The University of Tokyo	Toshio Koike
1.3 Development of Advanced Data Visualization and Mining System	Center for Spatial Information Science, The University of Tokyo	Eiji Ikoma
1.4 Development of advanced network infrastructure for huge scale data collection and distribution	Information Technology Center, The University of Tokyo	Masaya Nakayama
1.5 Development of data storage and fusion framework for active data sensor network	National Agriculture and Bio-oriented Research Organization	Seishi Ninomiya
2. Development of Mutual Use and Information Service System		
2.1 Development of Ontology Registry (Recording, Comparing, Editing System of Ontology Information)	Center for Spatial Information Science, The University of Tokyo	Ryosuke Shibasaki
2.2 Information Sharing and Access Support Service using Ontology Registry	Center for Spatial Information Science, The University of Tokyo	Ryosuke Shibasaki
2.3 Distributed data fusion system in agriculture	National Agriculture and Bio-oriented Research Organization	Takuji Kiura
2.4 Development of Distributed Satellite Data Integration System	Japan Aerospace Exploration Agency	Osamu Ochiai
3. Research Development on Earth Observation Data Collection and Data Quality Control		
3.1 Improvement of Satellite Data Collection and Data Processing Methods		
3.1.1 Development of Multidementional Data Integration for Satellite Data	School of Engineering, Shibaura Institute of Technology	Mikio Takagi
3.1.2 Development of satellite data reception, management and archiving system	Institute of Industrial Science, The University of Tokyo	Yoshifumi Yasuoka
3.2 Satellite Data Set processing	Japan Aerospace Exploration Agency	Chu Ishida
3.3 Integration of hydrometeorological data	School of Engineering, The University of Tokyo	Toshio Koike
3.4 Integration of Oceanographic Data	Ocean Research Institute, The University of Tokyo	Yutaka Michida
3.5 Integration of research in-situ data	School of Engineering, The University of Tokyo	Katsunori Tamagawa
4. Research and development on advanced information application technology		
4.1 The earth environmental data integration and information fusion for improvement of numerical weather forecasting	Numerical Prediction Division, Japan Meteorological Agency	Ko Koizumi
4.2 Application Research of Earth Observation Integrated Data for Water Management and Its Fused Information		
4.2.1 Development of Utilization System in Water Management Application	Ministry of Land, Infrastructure and Transport, National Institute for Land and Infrastructure Management	Kazunori Wada

4.2.2 Development of atmosphere, land and river channel coupled model	School of Engineering, The University of Tokyo	Toshio Koike
4.3 Utilization of integrated earth observation data and information fusion for optimal agricultural management		
4.3.1 Totalized data assimilation and integration methods for crop managements	National Agriculture and Bio-oriented Research Organization	Ninomiya Seishi
4.3.2 Soil-Vegetation Model and Its Model Integration	Graduate School of Agricultural and Life Sciences, The University of Tokyo	Masari Mizoguchi,
4.4 Utilization of integrated earth observation data and information fusion for Advance Climate Information		
4.4.1 Study on the verification and application of JMA Climate Data Assimilation System	Global Environment and Marine Department /Climate Prediction Division, Japan Meteorological Agency	Tomoaki Ose
4.4.2 Analysis of earth's global warming phenomenon with use of the integrated database	Center for Climate System Research, The University of Tokyo	Teruyuki Nakajima