

Session scope, and hosting ICSU-WDS International Program Office in Japan

MURAYAMA, Yasuhiro^{1*}

¹National Institute of Information and Communications Technology

Data in Earth, planetary and space sciences is growing in size explosively, is becoming heterogeneous in nature, and is requiring multidisciplinary interactions to related scientific fields and more general society. As the importance of the data is increasingly stressed, further efforts to use and publish such data in various ways are expected for communications with general society. Database or data center works in individual institutes have become more active; at the same time discussions to make interdisciplinary cooperation or fusion between databases, organizations, and data systems are strongly required. While international unions, programmes, and bodies such as IUGG, IPY, IRDR, WMO, UNESCO, etc., have been launching their own data activities and/or are showing the interest in data issues separately, they are now discussing or approaching collaboration with related activities. A new programme, ICSU World Data System (WDS), was launched in 2008, where one of the most important goals is to realize "system of data systems". In November 2010 ICSU decided to establish its international programme office (IPO) in Japan. This session welcomes reports and discussions on domestic and international data activities, and development of system of data systems and key technologies for international data sciences and data systems, as well as it fosters and facilitates various exchanges and mutual understanding between multidisciplinary science data activities/programmes/institutes; also this session welcomes discussions on activities and communities in Japan and how they can interact with ICSU-WDS and other related activities to aim at the shared interests and goals.

Keywords: data system, multidisciplinary science data, interoperability, Science Council of Japan, WDS, International Council of Science

ICSU World Data System-Global Data for Global Science

MOKRANE, Mustapha^{1*}

¹ICSU World Data System-International Programme Office

Perhaps the most compelling arguments for the long-term curation of scientific data sets is embodied in the current debate about the "anthropocene," and the extent to which human activity changes our planet. The time scales associated with such change are of course central to the discussion, and can only be ascertained if continuous time series of sufficient duration are collected and maintained. Consequently, scientific research relevant to that debate is increasingly data driven: from data assimilation to long-term time series and beyond. It is now well established that data have an intrinsic value that outlast current science foci.

Although new information and communication technologies encourage innovation and permit individual scientists and institutions to make data and information easily available, the web is constantly changing and somewhat chaotic. URLs disappear and previously available information can be lost without trace overnight. Data can be managed by individuals or groups in voluntary distributed systems on the internet but quality assurance and long-term accessibility issues are frequently neglected. For instance, reliable and systematic migration of data holdings to new storage technologies is often beyond the resources of all but the best supported data repositories. Thus, data sets collected only every few decades-for instance, data from the International Polar Year-are potentially at risk, unless concerted efforts are made to guarantee their long-term, sustainable curation.

A primary goal of the ICSU World Data System is to foster such efforts, and to support the long-term ICSU vision for a world in which science is used for the benefit of all, with universal and equitable access to high-quality scientific data. This, we argue, is closely linked both to scientific progress and technological advances, and calls for a fresh view of the concept of "publishing" carefully vetted data sets in trustworthy repositories with long-term sustainability prospects.

Keywords: Data management, Data stewardship, Data repositories, Data publication, Open access to scientific data



Contribution of Data Community of Japan to ICSU World Data System

WATANABE, Takashi^{1*}

¹Solar-Terrestrial Environment Laboratory, Nagoya University

Current activities of the national committee of the ICSU World Data System (WDS) of the Science Council of Japan are reported. The First ICSU WDS Conference "Global Data for Global Sciences" was held in Kyoto on 3 - 6 September, 2011. Around 155 participants (including 86 local participants) from over 22 countries attended. Participants included representatives of data centers and data services covering a wide range of scientific disciplines, data scientists and engineers working in a variety of fields such as natural sciences, social sciences and information technologies, as well as data publishers. The 23 invited talks, 36 contributed talks, over 70 poster papers, and 5 exhibits enabled the nascent WDS community to engage in effective scientific collaboration and provided a constructive forum for lively exchanges of views and ideas. Important feedback was also provided to the WDS Scientific Committee during an open forum, that will certainly influence and help shape the World Data System in the future. The Proceedings of the conference will be published as a special issue of the Data Science Journal of CODATA. Another important WDS-related movement in Japan is establishment of the WDS International Program Office (WDS-IPO) in 2012. This office is hosted by the National Institute of Information and Communications Technology (NICT). After these activities, the WDS community in Japan has a plan to establish a network of data centers to enforce Japanese contribution to the WDS.

Keywords: Data, ICSU, WDS, Japan

U01-04

Room:105

Time:May 23 09:45-10:05

Scientific Information Commons and World Data System (tentative)

IWATA, Shuichi^{1*}

¹The University of Tokyo

Under preparation

Data Integration and Information Fusion towards the Integrated Human Security

KOIKE, Toshio^{1*}

¹The University of Tokyo

To achieve Integrated Human Security, including the security of water, food, energy, health and ecosystem services, nations first need to share comprehensive and accurate data and information, then prepare various measures to prepare for threats and disasters in advance of their occurrence, provide society with timely support and sound decision making, and establish trans-boundary safety networks towards a resilient society. We need data integration infrastructure which enables scientists, practitioners, decision-makers, citizens and other stakeholders to work together toward *end-to-end* cooperation.

To promote effective multi-sectoral, interdisciplinary collaboration based on coordinated and integrated efforts, the Global Earth Observation System of Systems (GEOSS) is now developing a "GEOSS *Water Cycle Integrator (WCI)*", which integrates "Earth observations", "modeling", "data and information", "management systems" and "education systems". GEOSS/WCI sets up "work benches" by which partners can share data, information and applications in an interoperable way, exchange knowledge and experiences, deepen mutual understanding and work together effectively to ultimately respond to issues of both mitigation and adaptation. GEOSS/WCI enhances the coordination of efforts to strengthen individual, institutional and infrastructure capacities, especially for effective interdisciplinary coordination and integration.

Inter-University Upper Atmosphere Global Observation Network (IUGONET)

TSUDA, Toshitaka^{1*}, SATO, Natsuo², FUJII, Ryoichi³, ONO, Takayuki⁴, YUMOTO, Kiyohumi⁵, IYEMORI, Toshihiko⁶, SHIBATA, Kazunari⁷, HAYASHI, Hiroo¹, HORI, Tomoaki⁸, TANAKA, Yoshimasa², KOYAMA, Yukinobu⁶, ABE, Shuji⁵, SHINBORI, Atsuki¹, UMEMURA, Norio⁸, YONEDA, Mizuki⁹, UENO, Satoru⁷, KANEDA, Naoki⁷

¹RISH, Kyoto University, ²National Institute of Polar Research, ³Nagoya University, ⁴Division of Geophysics, Graduate School of Science, Tohoku University, ⁵SERC, Kyushu University, ⁶DACGSM, Graduate School of Science, Kyoto University, ⁷Kwasan Observatory, Faculty of Science, Kyoto University, ⁸STEL, Nagoya University, ⁹PPARC, Tohoku University

IUGONET is a joint project aiming at establishment of a data exchange system for the Earth's upper atmosphere observations. The participating members are the National Institute of Polar Research (NIPR), Tohoku University, Nagoya University, Kyoto University, and Kyushu University. We have built a metadata database (MDB) of ground-based observations that have been continued by means of a global network of radars, magnetometers, optical sensors, helioscopes, and so on. MDB provides contacts and basic information about the observed data. We intend to provide researchers with a seamless data environment linking databases spread across the member institutions. This MDB will be of great help in conducting comprehensive analyses with various observational data to clarify the mechanisms of the long-term variations in the upper atmosphere, which may be affected by global warming, solar activities, etc. In particular, IUGONET will greatly contribute to CAWSES (Climate and Weather of the Sun- Earth System), which is an international collaborative program promoted by SCOSTEP.

Keywords: CAWSES, metadata, upper atmosphere

Framework for International Collaborations in Planetary Data Archives

YAMAMOTO, Yukio^{1*}, SHINOHARA, Iku¹, HIRATA, Naru², KASABA, Yasumasa³

¹Japan Aerospace and Exploration Agency, ²University of Aizu, ³Tohoku University

The framework of sharing and utilizing planetary data archive has just started since early in 2000's. Historically, NASA has been facilitated the data archives along with planetary explorations, Planetary Data System (PDS). The PDS is sophisticated but it is difficult to implement for other countries due to demandingness. While NASA is promoting PDS as an archive system, ESA has developed its own data archiving system, Planetary Science Archive (PSA). The format of PSA and PDS is almost same, but the review process for documents is different. NASA and ESA discussed about the differences between PSA and PDS and considered to make a new standard. At the same time, other countries such as Japan, China, India, and etc. has started the Moon Race, and expected to obtain a huge dataset of the Moon. The activities of NASA and ESA extended to the international collaborations, and International Planetary Data Alliance, IPDA, was established. Japanese members join the IPDA and discuss about the new framework of how to share the planetary data archives.

Keywords: IPDA, PDS, PSA

International and Interdisciplinary Data Access: the IPY Experience

CARLSON, David^{1*}

¹IPY International Programme Office

Consistent with existing international guidelines, the International Polar Year 2007-2008 (IPY) adopted a free and open data access policy. This policy received wide agreement from leaders of IPY Projects as well as endorsement from national data centres and from international sponsoring organizations ICSU and WMO. However, implementing and supporting a free and open access policy across the breadth of IPY science and across the range of participating data centres proved difficult. The barriers and challenges encountered during IPY will seem familiar to many World Data Centres, while some of the IPY solutions, including a Polar Information Commons and renewed attention to data publication, offer new possibilities and new insights into practical aspects of open data access. These lessons from IPY should inform the goals and practices of the new ICSU World Data System initiative.

Keywords: Data access, International Polar Year, Polar Information Commons, Data publication

THE STATE OF IPY DATA: JAPANESE CONTRIBUTION AND LEGACY

KANAO, Masaki^{1*}, KADOKURA, Akira¹, OKADA, Masaki¹, YAMANOUCHI, Takashi¹

¹National Institute of Polar Research

Diverse data accumulated by many science projects make up the most significant legacy of the International Polar Year (IPY2007-2008). The Polar Data Center (PDC) of the National Institute of Polar Research (NIPR) has a responsibility to manage these data for Japan as a National Antarctic Data Center (NADC) and as the World Data Center (WDC) for Aurora. During IPY, a significant number of multidisciplinary metadata records have been compiled from IPY- endorsed projects with Japanese activity. A tight collaboration has been established between the Global Change Master Directory (GCMD), the Polar Information Commons (PIC), and the newly established World Data System (WDS).

The status of IPY data-management in Japan has been summarized in this presentation. Many dedicated data service tasks have been conducted by the staffs of PDC in NIPR as a member of NADC under SCAR. Several different aspects of scientific data collected in the polar region have great significance for global environmental research in this century. To construct an effective framework for long-term strategy of the polar data, data must be made available promptly and new Internet technologies such a repository network service like the PIC must be employed.

In addition to the activities in polar science communities of SCAR and the International Arctic Science Committee (IASC), tighter linkages must be established with other cross-cutting science bodies under ICSU, such as CODATA, and WDS. Linkages among these data-management bodies need to be strengthened in the post IPY era.

Keywords: International Polar Year, National Antarctic Data Center, Data Management, Metadata Portals, Polar Information Commons, World Data System

Toward a new phase of Arctic research data activity

YABUKI, Hironori^{2*}, KAWAMOTO, Haruko¹

¹National Institute of Polar Research, ²Japan Agency for Marine-Earth Science and Technology

In Arctic region, various changes are being caused by global warming. In previous studies, the decrease of Arctic sea ice extent, the increase in soil temperature in the region of Siberia, permafrost melting, the increase in Arctic river runoff, reduction in snow cover has been revealed. Also the impact of human activities and ecosystems due to these changes is concerned. The actual condition and mechanism of environmental change in arctic has not been elucidated. Previous studies have been carried out to separate the atmosphere, oceans, by land. Arctic is a system consisting of atmosphere, ocean, land surface, from snow and ice, these systems including the phenomenon of different spatial scales and time scale, respectively. In order to clarify the variability of the Arctic environment, through interdisciplinary research, research is needed in using a database that integrates the results of observation and research data across multiple areas.

In this presentation, we introduce the direction of data using and archiving in the National Institute of Polar Research is a research project focused on the Arctic region.

Keywords: Arctic, Environment, Global Warming

Activities on the development of GEOSS and GEO Grid

IWAO, Koki^{1*}

¹National Institute of Advanced Industrial Science and Technology

The Global Earth Observation System of Systems (GEOSS) is an intergovernmental initiative under the intergovernmental Group on Earth Observations (GEO). GEOSS will proactively link together existing and planned Earth observing systems around the world, and support the development of new systems where gaps currently exist. GEOSS aims to provide decision-support tools to a wide variety of users. This system of systems will promote common technical standards so that data from the thousands of different instruments can be combined into coherent data sets. The GEO Grid, which is the AIST initiative. Its concept and goal are similar to that of GEOSS. In the presentation, these activities and the relationships will be introduced.

Keywords: Global Earth Observation System of Systems, Group on Earth Observations, GEO Grid

U01-12

Room:105

Time:May 23 14:25-14:55

Integration of heterogeneous data sources of Russian-Ukrainian WDS Segment based on ontology and agent-oriented approach

YEFREMOV, Kostiantyn^{1*}

¹World Data Center for Geoinformatics and Sustainable Development, ²National Technical University of Ukraine "Kyiv Polytechnic Institute"

temporary abstract text

Keywords: World Data System, ontology, multiagent system, data sources integration, various nature data

Application Layer in Science Data Systems: Case Study of "Digital Typhoon" and "2011 Great Tohoku Earthquake"

KITAMOTO, Asanobu^{1*}

¹National Institute of Informatics, ²JST

This paper discusses methodology for the effective usage of science data systems using "application layer" as the keyword. The data itself is usually just a collection of numbers, and this is useful only for people who share the context of data such as how to read, process, and interpret numbers. The problem of context is critical not only for the general public, but also for researchers working in other domains. We claim that a science data system with improved accessibility and usability should have a well-designed application layer which offers context to read, process, and interpret data. We introduce two of our projects, namely "Digital Typhoon" and "2011 Great Tohoku Earthquake," to discuss how the application layer can be designed from the viewpoint of information architecture, data integration, information visualization, and social media.

Keywords: science data system, application layer, information architecture, data integration, information visualization, social media

Toward a Big Data Science: A challenge of Science Cloud

MURATA, Ken T.^{1*}, WATARI, Shinichi¹, NAGATSUMA, Tsutomu¹, KUNITAKE, Manabu¹, WATANABE, Hidenobu¹, YAMAMOTO, Kazunori¹, KUBOTA, Yasubumi¹, MURAYAMA, Yasuhiro¹, KATO, Hisao¹, TSUGAWA, Takuya¹, SHINAGAWA, Hiroyuki¹, JIN, Hidekatsu¹, TANAKA, Takashi², SAKAGUCHI, Kaori¹, SAITO, Shinji¹, NISHIOKA, Michi¹, ISHIBASHI, Hiromitsu¹

¹NICT, ²Kyushu University

During these 50 years, along with appearance and development of high-performance computers (and super-computers), numerical simulation is considered to be a third methodology for science, following theoretical (first) and experimental and/or observational (second) approaches. The variety of data yielded by the second approaches has been getting more and more. It is due to the progress of technologies of experiments and observations. The amount of the data generated by the third methodologies has been getting larger and larger. It is because of tremendous development and programming techniques of super computers.

Most of the data files created by both experiments/observations and numerical simulations are saved in digital formats and analyzed on computers. The researchers (domain experts) are interested in not only how to make experiments and/or observations or perform numerical simulations, but what information (new findings) to extract from the data. However, data does not usually tell anything about the science; sciences are implicitly hidden in the data. Researchers have to extract information to find new sciences from the data files. This is a basic concept of data intensive (data oriented) science for Big Data.

As the scales of experiments and/or observations and numerical simulations get larger, new techniques and facilities are required to extract information from a large amount of data files. The technique is called as informatics as a fourth methodology for new sciences.

Any methodologies must work on their facilities: for example, space environment are observed via spacecraft and numerical simulations are performed on super-computers, respectively in space science. The facility of the informatics, which deals with large-scale data, is a computational cloud system for science.

This paper is to propose a cloud system for informatics, which has been developed at NICT (National Institute of Information and Communications Technology), Japan. The NICT science cloud, we named as OneSpaceNet (OSN), is the first open cloud system for scientists who are going to carry out their informatics for their own science.

The science cloud is not for simple uses. Many functions are expected to the science cloud; such as data standardization, data collection and crawling, large and distributed data storage system, security and reliability, database and meta-database, data stewardship, long-term data preservation, data rescue and preservation, data mining, parallel processing, data publication and provision, semantic web, 3D and 4D visualization, out-reach and in-reach, and capacity buildings.

Figure is a schematic picture of the NICT science cloud. Both types of data from observation and simulation are stored in the storage system in the science cloud. It should be noted that there are two types of data in observation. One is from archive site out of the cloud: this is a data to be downloaded through the Internet to the cloud. The other one is data from the equipment directly connected to the science cloud. They are often called as sensor clouds.

In the present talk, we first introduce the NICT science cloud. We next demonstrate the efficiency of the science cloud, showing several scientific results which we achieved with this cloud system. Through the discussions and demonstrations, the potential performance of sciences cloud will be revealed for any research fields.

Keywords: Big Data, Science Cloud, OneSpaceNet

Access and scientific exploitation of planetary plasma datasets with the CDPP/AMDA web-based facility

ANDRE, Nicolas^{1*}, the CDPP Team¹, the Europlanet-RI IDIS plasma node (IWF, Graz and CDPP, Toulouse)¹

¹IRAP, CNRS/Universite Paul Sabatier, Toulouse, France

The field of planetary sciences has greatly expanded in recent years with space missions orbiting around most of the planets of our Solar System. The growing amount and wealth of data available make it difficult for scientists to exploit data coming from many sources that can initially be heterogeneous in their organization, description and format. It is an important objective of the Europlanet-RI and IMPEx projects (supported by EU within FP7) to add value to space missions by significantly contributing to the effective scientific exploitation of collected data; to enable space researchers to take full advantage of the potential value of data sets. To this end and to enhance the science return from space missions, innovative tools have to be developed and offered to the community. AMDA (Automated Multi-Dataset Analysis, <http://cdpp-amda.cesr.fr/>) is a web-based facility developed at CDPP Toulouse in France (<http://cdpp.cesr.fr/>) for on line analysis of space physics data (heliosphere, magnetospheres, planetary environments) coming from either its local database or distant ones. AMDA has been recently integrated as a service to the scientific community for the Plasma Physics thematic node of the Europlanet-RI IDIS (Integrated and Distributed Information Service, <http://www.europlanet-idis.fi/>) activities, in close cooperation with IWF Graz (<http://europlanet-plasmanode.oeaw.ac.at/index.php?id=9>). We will report the status of our current technical and scientific efforts to integrate in the local database of AMDA various planetary plasma datasets (at Mercury, Venus, Mars, Earth and moon, Jupiter, Saturn) from heterogeneous sources, including NASA/Planetary Data System (<http://ppi.pds.nasa.gov/>). We will also present our prototype Virtual Observatory activities to connect the AMDA tool to the IVOA Aladin astrophysical tool to enable pluridisciplinary studies of giant planet auroral emissions.

Keywords: planetary plasma, data archive, virtual observatory, tool, access, conditional search

Substorm Zoo - a browser-based tool for space weather research and teaching

TANSKANEN, Eija^{1*}

¹Finnish Meteorological Institute, Helsinki, Finland

Large amount of high-resolution measurements are nowadays available from different heliospheric locations. It has become an issue how to best handle the ever-increasing amount of information about the near-Earth space weather conditions, and how to enable the social data analysis. To resolve the problem, we have developed an interactive web interface, called Substorm Zoo (www.substormzoo.org), which we expect to become a powerful tool for scientists and a useful tool for public. The aim is to (1) provide a combined data repository for different heliospheric measurements including the geomagnetic activity indices with a possibility to customized views, (2) enable the use of pre-identified event lists, creation and sharing of own lists, (3) allows discussion on individual activity events e.g. substorms from the users of the site, and (4) enable the interactive data analysis on-line with a possibility to write and share comments. In this paper, we will present the basic features of Substorm Zoo and give examples of the use for educational, scientific and public outreach purposes.

Keywords: Interactive web interface, Tool for data analysis, Space weather, Substorms, Event lists