

Brief of each lecture

Toward a development of “Knowledge Infrastructure”

Michiko Kawanabe (Director of Business, Science and Technology Division, NDL)

The NDL collects and preserves publications from inside and outside Japan as its mission. Recently we pay more and more attention to collecting electronic publications such as web resources based on our objectives.

The NDL formulated the “Third basic plan for the development of science and technology information” in March 2011. The third basic plan indicates that the NDL will contribute to actively building “Knowledge Infrastructure” which aims to make a circulation system for preserving, sharing and reusing various kinds of information including data. This is why we are hosting this symposium today. We hope the symposium will be useful for the future of libraries.

In addition, the outline of the symposium will be introduced at this speech.

【Lecture】 Recent activities in global research data management

Dr. Yasuhiro Murayama

(Director of the Integrated Science Data System Research Laboratory, National Institute of Information and Communications Technology/ Visiting professor, Research Institute for Sustainable Humanosphere, Kyoto University)

Open data is not only the subject discussed in the last G8 meeting 2013, but also can be a wide-spread argument and can be a substantial point of conducting the science. In the modern science scheme, results of scientific research can be validated by, e.g., its reproduction or statistical significance particularly in some natural science areas such as physics or so. However recently scientific results are found not to be necessarily reproducible. Today's society has increasingly a big concern with climate change and huge earthquake etc., where scientific research results may directly affect political decision making. Validation of scientific papers will affect mutual trust between science and society. Here electronic data which can be linked to scientific papers in data citation scheme, will be part of evidence of our scientific truth. So library in future could play really important role in repository work for both data and literature, where the library can be indispensable component to construct the science-society mutual trust; in other words, the system of Science for society will require a large number of stakeholders from scientists and researchers in universities, research institutes, and companies, to data curator, publishers of papers and data, and the library for the electronic papers and data. In comparison to the hundred-year history of printing culture for the scholarly communications, Internet and hard disk drives have only tens-of-years history. Human beings are now challenging this new system of electronic way to conduct the right science with society in the right direction. Data management works like ICSU-WDS from academic activity, and RDA related to governmental arrangement are part of such international challenges of the international community. We need careful discussions to promote those activities, but with bright hope to the human society with the beautiful infrastructure called “science”.



【Keynote Lecture】 TIB's action for research data management as a national library's strategy

Dr. Peter Loewe

(Head of Development, German National Library of Science and Technology/ Visiting Scientist [High Performance Computing], GFZ German Research Centre for Geosciences)

This lecture will give an overview over the approach and actions of the German National Library of Science and Technology towards a sustainable research data management strategy for the next decades.

The German National Library of Science and Technology (German: Technische Informationsbibliothek (TIB)), focuses on the fields of engineering, technology, and the natural sciences. Founded in 1959, the library operates in conjunction with the Leibniz Universität Hannover. It is jointly funded by the Federal Ministry of Education and Research (BMBWF) and the 16 German states, having an annual budget of 14.7 million Euro and a staff of about 400 people in 2012. The collection of the TIB contains 9 million items and 52,700 journals in 2012. This makes it currently the largest science and technology library in the world. It serves customers worldwide and is part of national, European and global networks.

In addition to acquiring scientific literature, TIB also conducts applied research in areas such as the handling of non-textual materials and data visualization. For this, a new research and development department was established in 2013.

Finding solutions to cope with the current divide separating scientific publications and the underlying research data is a major challenge for TIB. Potential losses of scientific achievements must be prevented by making information traceable, citable and accessible. Part of a solution strategy must be global access to data sets and metadata through catalogues, the monitoring of emerging technology trends and the use of persistent identifiers for data such as Digital Object Identifier (DOI). In 2005, the TIB became a DOI registration agency for primary data. This was transferred to a new worldwide DOI agency, named DataCite, in 2009. The global DataCite consortium supports researchers, data centers and publishers by providing standards, workflows and best-practices for the use of DOI for their referencing and publication. DataCite consists currently of 19 full members and 10 affiliated members in North America, Europe, Asia and Australia. By 2014, over 2,500,000 DOI were registered by 272 data centers.

Within Germany, TIB shares its application-oriented research via the GOPORTIS network, which was founded in 2009. GOPORTIS is a strategic network of the three German national libraries, including the National Library of Medicine and the National Library of Economics. Cooperation focuses on the provision of scientific content, research and innovation and political work.

In the RADAR project, the TIB is reaching out to the “small sciences” and the industries to establish research data infrastructures to facilitate research data management. RADAR makes a key contribution to ensure a better availability, sustainable preservation and publishability of independent research data. The RADAR consortium consists of the Leibniz Institute for Information Infrastructure (FIZ), the Karlsruhe Institute of Technology (KIT-SCC), the Leibniz Institute for Plant Biochemistry (IPB), the Ludwig-Maximilians-Universität München (LMU) and the TIB. As a generic, interdisciplinary service, RADAR pursues a two-stage approach with a non



subject-specific starter package for preserving research data in the context of good scientific practice and a superior package for preserving data with integrated data publication. Both packages will be part of a business model which includes one-off payments depending on data volumes and storage periods. Using this cost model, the project aims to establish a sustainable operation environment for the data archive. A close cooperation between scientists, expert organisations and publishing houses will ensure that the RADAR infrastructure will be developed according to the needs of the future research landscape.

Regarding the role of research libraries in the new era of Big Data and Data Science, it is imperative to consider how research data management might evolve in the next decades and how to take on upcoming chances and threats. For the European Union, the Final report of the High Level Expert Group on Scientific Data, titled “Riding the Wave”, develops scenarios for the relationship of science towards data management, the citizen, data sets, students and incentives for data sharing. A set of milestones is staked out until the year 2030.

For Germany, the project “Radieschen: Requirements for a multi-disciplinary research data structure”, conducted by GFZ Potsdam, describes five future scenarios until and beyond the year 2020. Based on community polls in Germany and Europe, the scenarios describe possible interaction patterns between scientists, libraries, computation centres and scientific publishers.

Considering the scenarios for Europe and Germany, and the power of disruptive innovations, a promising strategy for the provision of research infrastructures could be a modularized service portfolio, based on a common platform. This would enable the stakeholders to adapt the services flexibly according to the changing requirements of science, while allowing for the long term evolving of the underlying platform. This will bridge the gap between infrastructure’s need for stability while allowing for the required flexible, yet potentially short-lived applications for science.

【Case Study 1】 State of data sharing in Agricultural Research

Mr. Takuji Kiura

(Senior Researcher, Agriinformatics Division, Agricultural Research Center, National Agriculture and Food Research Organization)

Agricultural researches are multidisciplinary, and require variety of data in wide areas relating Agriculture. In agricultural research, the environment can not be controlled, i.e. re-search data are depending the location and time, and cannot be observed, measured, or obtained again. So data sharing and preservation are important. Many organization, FAO, MAFF AFRIT, Independent Administrative Organizations under MAFF (NIAES, NIAS, NARO, etc), Universities are sharing data. But unfortunately it is difficult to find valuable data and related data, and the use them. Other agricultural research data still be unshared even if Japanese government financially supports those researches. Data Integration & Analysis System (DIAS) provides data archiving, sharing, integration, and analysis functions for earth observation data, and is used for many research field. DIAS for agricultural research data is expected.



【Case Study 2】 Data Preservation and Dissemination at SSJDA (Social Science Japan Data Archive)

Dr. Hiroki Sato

(Professor, Interfaculty Initiative in Information Studies, Institute of Social Science, The University of Tokyo)

Whilst the basic function of data archives is to collect, store, and provide micro-data, a more important function is to provide researchers with an opportunity to 'reproduce' analysis of past empirical research through secondary analysis. Japan produces a large number of social surveys, which are carried out not only by academics but also by various organisations such as newspaper companies and government agencies.

Despite the country's being a 'social survey giant', data often disappeared after the publication of papers and reports. Until recently, there was no system which acquired, curated, and stored data, or provided data for secondary analysis. This lecture illustrates the social importance of depositing data for those who wish to conduct secondary analysis, and the role of data archives, focusing in particular on the aims and goals of the Social Sciences Japan Data Archive (SSJDA) managed by the Center for Social Research and Data Archives, Institute for Social Science, University of Tokyo.

