

## ICSTI 2008 Public Conference

---

# New Frontiers for Scientific and Technical Information

*CODEX International Hotel, Samsung-dong, Seoul, Korea  
June 11–12, 2008*

David J. Brown  
*Director, SCR Publishing Ltd.*

### **Background**

The International Council for Scientific and Technical Information (ICSTI) held its 2008 Public Conference in Seoul, Korea from 11th to 12th June, 2008. This was a departure from recent practice – it has been more typical for ICSTI public conferences to be held in Europe and North America. The decision to hold the 2008 Summer Conference in Korea was fully vindicated by the results. It was a very well organised event for which the local host – Korea Institute of Science and Technology Information (KISTI) – can claim full credit. Attendance was well over three hundred, some two to three times the usual ICSTI numbers at such events. Each attendee received a comprehensive programme guide and proceedings which proved invaluable as support material for the presentations. Leading experts from across the world shared their views and experiences with the audience. Included within the programme was the formal international launch of a major new initiative sponsored by ICSTI. Overall the logistics went very smoothly and professional staff ensured that the conference at the CODEX International Hotel, and the related social programmes, were a great success.

This was a perfect template for a revitalised ICSTI and sets a standard for future conferences including the 2009 to be held under CISTI's auspices in Canada. Everyone involved in the Seoul 2008 Public Conference, particularly the local hosts KISTI, are to be congratulated for their organisation and attendance to detail.

### **Table of contents for the ICSTI 2008 Public Conference**

#### Opening Ceremony

Byeong-Tae Yang, President KISTI, Korea  
Sang-Wan Han, Chairman, Korean Library and Information Policy bureau  
Jong-Koo Park, Deputy Minister, MEST, Korea

Hee-Yul You, Chairman, KRCFST, Korea  
Herbert Gruttemeier, President, ICSTI, France

#### General Sessions

New horizons for academic communication, Prof. Dong-pil Min, Seoul, Korea  
The Finnish information environment and future challenges, Kirsi Tuominen, VTT, Finland  
Enhancing the use of Scientific and Technical Information, Young-Mee Chung, Korea  
Information service strategy for STM researchers in the web age, Richard Boulderstone, BL, UK

#### Session 1 – New trends in scholarly communication

Trends in secondary publishing, Helle Lauridsen, ProQuest  
Globalisation of Korean medical journals, Prof. Choon-Shil Lee, Korea

#### Session 2 – Track 1 – STI: Dynamic engines for R&D

R&D Navigator: National Technical Information System, Inseok Song, KISTI, Korea  
Library as a partner for research, Prof. Toshiro Minami, KIIS, Korea  
Integration of knowledge organisation systems, Rudzhero Gilyarevskiy, VINITI, Russia  
WorldWide Science.org – Accelerating global scientific discovery, Walter Warnick, USA

#### Session 3 – Track 2 – Scientific data management and integration

Scientific data management in the global e-science environment, Kum Won Cho, KISTI, Korea  
Research Information Centre – a virtual research environment, Alex Wade, MSN, USA and Lee Ann Coleman, BL, UK  
Connecting biomolecular databases and the literature, Peter Stoehr, EBI, UK  
Biological diversity information flow in KBIF, Hyung-Seon Park, KISTI, Korea

#### Track 3 – Special Session on Licensing issues

SUSHI: A beginner's guide to NISO's standard, Oliver Pesch, EBSCO, USA  
KESLI – Science Direct proposal from Elsevier Korea  
KESLI – Springer proposal for next year from Springer Korea

#### Session 4 – Track 1 – STI service innovation in East Asia

S&T information service and support system in China, Jiachang Chen, ISTIC, China  
Access to and dissemination of STI in Japan, Yukiko Sone, JST, Japan  
Innovation of S&T information services in Korea, Tae-Sul Seo, KISTI, Korea

#### Session 5 – Track 2 – Scientific data management and integration

Scientific data management at Pohang Light Source, In Soo Ko, PAL, Korea  
Factual databases developed by KISTI, Sang-Ho Lee, KISTI, Korea  
DOI registration for scientific primary data, Jan Brase, TIB, Germany

#### Track 3 – Special Session on Evolution in Information Services in Portals

KESLI-WoS Proposal Presentation for next year, by Thomson Korea  
Google Scholar, Google, Korea  
Knowledge Search, present and future, Beom-Seok Seo, NHN, Korea

Session 6 – Role of the Library in Scholarly Communication in the Web Age, Jay Jordan, OCLC, USA  
Special Session – WorldWideScience Alliance signing ceremony, featuring Dr. Jeffrey Salmon, DoE, USA

## Introduction

The opening ceremony on Wednesday 11th June included short presentations by leading authorities in the Korean information space. The first speaker was **Byeong-Tae Yang**, President of KISTI in Korea who welcomed the delegates particularly from around the world to the Public Conference. Mr. Yang referred to the topic of the conference – identifying the New Frontiers for Scientific and Technical Information – by suggesting that the debate about global scholarly communication has currently become an issue involving considerable focus of attention. It transcends national frontiers. It has become international in scope with the issue of ‘open science’ being more in evidence than ever before. Nevertheless, despite the progress achieved so far continuous development of the scholarly and research communication processes are still necessary. In particular, information and knowledge needs to be shared across all global sectors. As such the speaker welcomed the many experts from across the world who were attending the ICSTI Public Conference to discuss the many New Frontiers for Scientific and Technical Information (STI). Korea intends to assume its part in ensuring increasing productivity in this area and to see this shared by all.

The next speaker was **Sang-Wan Han**, who offered his congratulations to the organisers on behalf of the Library and Information Policy Bureau in Korea of which he is chairman. Besides thanking the local organisers he also paid tribute to the many attendees who had travelled long distances to attend the event. He made the point that we are witnessing a paradigm shift in social and economic systems which were affected by, and affecting, the STI systems. This is in part a result of increased competition in the leverage of knowledge. This is as much evident in Korea as elsewhere. Information and Communications Technology (ICT) has become crucial in sustaining knowledge, its advancement and its application. Korea is noted for having the highest proportion of citizens accessing broadband in the world for example. Creating and sustaining an effective infrastructure is crucial.

The next speaker, who also offered his congratulations for the smooth organisation of the conference, was **Jong-Koo Park**. Mr. Park is the deputy minister at the Ministry of Education, Science and Technology (MEST) in Korea. The participation of such a high level official helped to stamp the conference with authority and visibility and which in itself may have helped to explain the attendance of many younger Korean academics in the audience, eager to learn about the official attitude to STI development in Korea. Mr. Park claimed there is a national drive to create social knowledge and scientific and technical information which the Korean government is willing to support. For this reason a new unit has been launched to coordinate this activity. Not only did the deputy minister make a plea for greater openness in the international sharing of information but he also made the point that Korea’s own development will depend on how it applies research and development (R&D) information in future.

The third speaker to congratulate the organisers was **Hee-Yul You**, chairman of the Korea Research Council of Fundamental Science and Technology. Mr. You agreed that the theme of the conference, on New Frontiers in STI, is particularly relevant and timely. We are seeing an explosion of information being generated by technologies such as the Sony Playstation 3, which is twice as fast as IBM’s traditional supercomputers. Two-way information systems are also becoming important – social interaction is taking place alongside the one-way flow of information dissemination which characterised the print-based publication model. The impact of such social collaboration and networking is yet to be fully felt.

So there was a consensus by these eminent speakers that the time was right to debate the new frontiers in STI and that the organisers had done a good job in bringing a stimulating programme together. **Herbert Gruttemeier**, president of ICSTI and member of the senior management team at INIST-CNRS in France, thanked the speakers and in particular thanked KISTI and the programme committee for

bringing together over 30 presentations around the particular conference theme. He indicated that there was a parallel in that ICSTI as an organisation had also undergone change in recent years, reflecting the changing nature of STI and the market for research information. ICSTI was not standing still, and an announcement of a major collaborative venture which endorses the aims of sharing information, outlined by the speakers, and supported by ICSTI, was due to be formally made later in the meeting. In conclusion, Mr. Gruttemeier thanked his colleagues at KISTI for all their hard work in enabling the ICSTI Public Conference to take place.

### **The General Sessions**

After the preliminaries were completed, the first speaker in the conference itself was **Dong-pil Min**, professor at the Basic Science Research Information Centre at the Seoul National University in Korea. Prof. Min pointed to the increasing trend whereby authors were depositing the results of their research experiments into their local institutional repositories, and the variety of benefits this conferred on the research process, for the academic societies and the benefits for the individual researchers themselves. He saw that this would lead to the formation of a world knowledge platform (Global Knowledge Platform or GKP) which will open up even greater benefits to society.

Prof. Min introduced the term 'Meson' in his presentation, which comes from the impact which advanced digital technology will have on the process of information dissemination. Meson is the technology of the medium which will facilitate improved, rapid dissemination as long as it can be divorced from its traditional reliance on providers and consumers of publications. In describing his thesis on the need for such improvements and the development of Global Knowledge Platforms the speaker made references to key information experts ranging from Schumpeter, through Attali to Harnad in supporting his premise that there is a need for more openness in global communication.

He took up the case made by the early speakers that there is a powerful case to be made for the rapid, widespread and free spreading of information, knowledge and ideas and that these three features are essential attributes of global knowledge platforms. There may be an information explosion taking place, but the crucial issue is to ensure that quality ideas rather than just information per se is disseminated effectively. This means that it is beholden on society to make sure that it is quality which rises to the top, and that these ideas are then distributed to those customers who can make most effective use of them. In this respect the medium for idea dissemination becomes as important as the content itself, a reference to McLuhan's claim in this area (where the latter suggested that the medium was in fact more important than content). Prof. Min compared the traditional schema of Producer–Customer–Content/Medium with his definition of Donar–Acceptor–Meson. Current technology enables the distinction between producer and customer to be accentuated – it enables the medium (meson) to be liberated from its traditional dependence on the provider (donar) and consumer (acceptor) of information.

Particular reference was made to work by Jacques Attali in his book 'A Short History of the Future'. The governing principle here is that there is a need for transparency. He then moved on to discuss the significance of ArXiv in the emerging system of digital communication. ArXiv has its roots in the exchange of printed preprints within the high energy physics community. Paul Ginsparg, in 1991, enabled an on-line archive of the electronic version of these preprints to be collected in one place and disseminated. There was an inbuilt auto-surveillance feature to this – the community itself retained governance over what publications were important and relevant. ArXiv became a model for a new paradigm – according to Ginsparg, 'It is the best way to communicate knowledge and hence create new knowledge'. Costs

of the publication system would be reduced by two orders of magnitude at least if an open publication system were to prevail.

The growing emergence of a Digital Network Environment produces a similar variety of new benefits to academia, research and society in general. Nevertheless, despite the many new digital developments, journal prices still continue to rise. To counter this mandates are being introduced by funding agencies and research centres such as, recently, Harvard's Faculty of Arts and Sciences which stipulate that research results need to be deposited in institutional repositories and made available openly through the Internet. Academic institutions need to confront what their role should be in such a new digital environment. They need to create, preserve and disseminate knowledge. They need to provide support for their staff and researchers to enable open dissemination of results. A Global Knowledge Platform is required – a networked consortium created by partnership which fosters trade without losing control in the process. An infrastructure should be put in place which enables efficiencies to be generated and this needs to take into account intellectual property (IP) issues and also to address issues of speed of access. We need to release ourselves from the obsession with IP and support the auto-surveillance (self-assessment) of publications to ensure quality is paramount and to promote the goals of the open and free Global Knowledge Platform.

The next speaker after the morning break was **Kirsi Tuominen**, head of Knowledge Solutions at VTT in Finland. Ms. Tuominen stressed the need for global collaboration between research centres to achieve progress in innovation. VTT is focused on industrial and applied research and has demonstrated its commitment to international collaboration through close ties being forged between Finland (VTT) and Korea (KISTI).

The speaker claimed that over the years there has been a migration from Resource-driven economic growth, through Investment-driven, to the current Innovation-driven economic growth. In future Wealth-driven growth will emerge. The strategies which Finland will adopt to parallel these trends will be included in a National Information Strategy which was due to be launched on 12th June 2008. As it stands, Finland is currently third behind Israel and Sweden in its R&D investment as a proportion of its gross national domestic product (gdp). Much of the Finnish effort in R&D effort is funded by corporations (Euros 4 billion). VTT funds Euros 217 million, Universities fund Euros 1 billion and Finnish ministries a further Euros 357 million. However, private investment dominates in the innovation stakes – 70% comes from private resources whereas only 30% are sourced from state coffers.

Ms. Tuominen reflected on the key assets which Finland has in promoting innovation. Firstly, there is a level of cooperation which exists stemming from Finland being more of a 'club' than a disparate country as far as R&D is concerned. People in the various parts of R&D in Finland know each other personally, which helps in stimulating research collaboration. It is strongly felt that, for the future, this emphasis on cooperation is important to enhance the overall national R&D effort. As tangible evidence that VTT does not just give lip service to this need for cooperation it was announced that VTT has established an office in Seoul (just the week before the conference) to facilitate cooperation between two highly innovative countries. Another major resource on which Finnish R&D can draw is the high level of research expertise which exists in Finland. Their scorings in HEFSA and WEF are invariably high. Finally, there is a strong public/private partnership with the public sector investment in innovation being an effective accelerator for private R&D and product development. There is also high level of trust, and transparency, in the information and knowledge creation processes in Finland.

Whilst investment in R&D in Finland is effective, it is not quite so good in the commercial exploitation of new products. This area needs improvement in Finland. Ease of submission and protection of

patent applications needs to be worked on, and better market research information for SMEs (small and medium-sized enterprises) needs to be provided in particular.

A new national Information Strategy in Finland is being unveiled which provides a public directive to support those services which require assistance for scientific and technical information provision. The new strategy will also enable a new Innovation University to be established in Helsinki. The Strategy will also support the creation of more national information hubs throughout the country. Though these are good measures the only pity is that they are a few years late in their establishment. But the aim of the new Information Strategy is to double the number of innovative organisations operating within Finland by 2015.

In line with the changing information world, VTT itself has also undergone change during the past 4–5 years. It has seen a migration from mainly basic research to more product development and an increased demand for useful market support services. It has begun to appreciate the value of ‘openness’ for Finnish innovation, and intends to exploit this in cooperation with other countries. This cooperation and partnership will involve both operational and strategic activities. They are encouraged by the position which the EU is taking on openness (open access) and collaboration. VTT has also largely disbanded its specific industry-focus, and now adopts a much more multidisciplinary approach to its research. VTT provides matrix level support services for the various research projects in such areas as strategic research, business solutions, joint ventures and other expert services. Each of VTT’s clients can take advantage of the central research services.

VTT’s staff of 2,740 (18% of whom are PhDs) are now mainly focused on applied research and product development. VTT operates from 10 cities within Finland, and with footholds in Brussels, St. Petersburg, Shanghai, in the UK, in Silicon Valley and now also on Seoul, Korea. The research projects are mainly organised into seven knowledge clusters and 46 knowledge areas. The main knowledge clusters include:

- Biotechnology
- Energy and forestry
- Microtechnology and sensors
- Telecommunications
- Digital Information Systems
- Industrial Systems
- Materials and Building

Only Business Solutions, a matrix service, will seek external contract research. Overall VTT’s commercial activities are stable at 232 million Euros (60% of research being self-financed). Ms. Tuominen represents VTT’s Knowledge Solutions matrix. This unit operates as far as possible utilising the latest informatics developments. Some 40 people are attached to the unit.

The next speaker was **Prof. Young Mee Chung**, from Yonsei University, Korea. The main theme of Prof. Chung’s presentation was to explore the linkages which exist between accesses to information in a digital world. This raises the whole issue of metrics, a rapidly expanding and topical field in the science of scientific information.

As the speaker pointed out, in essence there are two main metric areas – bibliometrics, which include Journal Citation Reports, co-citation reports, the h-index, etc. – and web metrics, which include download analyses. It is the analysis of web traffic which has become of increasing interest, and the National Scholarly Web Space (NSWS) for STI in Korea has become a centre for assessing the potential for measuring web usage patterns through link analysis. The premise is that web links would reflect the research productivity or the importance of a research institution or researcher in the same way that citation

counts reflect the importance of a scholarly journal, document or researcher. Link analysis is part of the PageRank algorithm invented and exploited so successfully by Google.

The NSWIS unit has collected online usage data from three major groups in Korea – from the universities (11 top ranking universities); from government sponsored research institutes (18, including KISTI) and from 17 national research information centres which are funded by the Korea Science and Engineering Foundation. Web crawlers were used to collect data on pages used and links made, on Web Impact Factors (WIF) and Web Use factors (WUF), from these centres during a one month period in the Spring of 2007. In total the database amounted to nearly 430,000 pages accessed and 150,000 links identified during this period from the above centres. The data was collected using two specialised web crawlers (SocScibot3 and RESPAC). The definitions used for pages, directories, domains and institutions were that of the Alternative Document Model proposed by Thellwall (2002), and each organisation had counts recorded according to these various categories.

The resulting data enables the behaviour patterns of users to be analysed. Patterns of linkages can illustrate the intensity of the various linkage types between institutions of a similar type (such as universities), and linkages between institutions with a different research orientation. From the extensive collection of data it became obvious that, in Korea, universities are the most active in online scholarly communication whereas the government-sponsored research institutes showed the least amount of linkages. Only a few of the national research institutes showed the same extent of linkages which the universities displayed.

In terms of subject areas, the life sciences and the natural sciences showed the greatest use of research information as indicated by the linkage pattern. Subjects with a moderate amount of information usage include agriculture, chemical engineering and medical science. Lowest use of linkages is within the chemistry, earth sciences, mathematics and information science.

Prof. Chung concluded by recommending that the database of linkages be increased through adding more institutions' traffic data so that the differences between institutions and disciplines be subject to more careful and detailed scrutiny. Establishing open access repositories would do much to provide more data and to facilitate more such inter-institutional linkages.

After the break for lunch, **Richard Boulderstone**, director of electronic Information Strategies (eIS) at the British Library, then pointed out that libraries are going through profound changes as they have been required to adapt to the digital world. We need to adapt in order to deal effectively with the growing amount of electronic information. In that respect Mr. Boulderstone described how the British Library had been changing its focus in recent years to remain as relevant today and tomorrow as it has been in the past. It has had to adapt to the fact that an increasing number of visitors and users of the British Library services are requiring digital access. This behavioural change, predicated on the basis of technological advances, needs to be understood.

To achieve this understanding a greater emphasis has been made on monitoring researchers' needs. A number of such studies have been made in the past including the recently published research into the 'Google Generation'. This has shown that those who have recently joined the ranks of research information searches are as confused about using the sophistication of search engines as their older colleagues. It also reflected on the amount of superficial 'power browsing' which goes on in the online environment, a clear difference from information seeking habits in the printed era. The British Library (BL) has also undertaken studies to determine what the real information needs are of the scientific and medical researcher operating at the frontiers of their subjects and disciplines. Support for UK scientific research has become a core strategic objective for the BL. This has been based in part on the extensive Science collection which is held at the Library.

The BL has also been active in developing an infrastructure to provide in perpetuity access to research material. The BL recognises that establishing partnerships in this area is really important. One particular theme which the speaker followed was the demographic nature of the changing market for scientific information. He identified the number of people undertaking R&D in the UK (200,000), the number of people in higher education, the undergraduates and postgraduates, the numbers availing themselves of document delivery services provided by the British Library and the number of onsite users. What is still missing is how many potential or latent users are not included in the above categories – how many knowledge workers are ‘disenfranchised’ by the current subscription and licensing based system. However, one successful development has been the creation of a centre at the BL which has its primary aim the support for small and medium-sized enterprises and individual entrepreneurs in developing their respective businesses.

Mr. Boulderstone also commented on the various recent changes in business models, and despite the heavy promotion given to the open access movement, this still represents less than 5% of all the (over 1 million) refereed research articles published each year. Besides research articles Mr. Boulderstone referred to the wide variety of other scientific information media held at the BL – technical standards, grey literature and reports, theses, official publications, sound recordings and web sites. In order to support the UK and global research community the BL has long provided access to this material through interlibrary loans, document delivery and more recently online delivery of material using a plug-in to ensure that the intellectual rights of the material’s owners are protected.

The BL has entered into partnerships with a number of organisations to push forward the UK national science programme. A new STM unit has been established to coordinate activities such as the creation of UK PubMed Central, to work with the US Department of Energy in the development of WorldWideScience.org (WWS.org) and collaboration with Microsoft on the Research Information Centre (RIC). Each of these involves a new departure for an organisation which has a strong legacy in the humanities and social sciences where electronic information dissemination has not been as evident as in STI. UKPMC involves transforming the US-based PubMed Central (PMC) system as created and maintained by the National Institutes of Health into something more relevant to the research process and ethos within the UK. WWS.org, another partnership activity, will be commented on later in the programme. RIC involves emulating the entire research process on the desk top, giving researchers immediate access to the full range of information requirements from the inception of the research idea to the publication of the final research paper, built around Microsoft’s Sharepoint platform. 18 different research teams are currently testing the initial phases of the RIC programme which covers the functions of discovery, funding, experimentation and dissemination, all through the same user interface and desktop system.

In addition, the British Library is fully committed to the creation of a Digital Library Programme which will enable all the research results, from all the relevant media, to be ingested, stored and made accessible – subject to relevant digital rights management being observed – to the research community. This is seen as a national requirement to ensure that an infrastructure is created to preserve scientific information in perpetuity. In this regard establishing partnerships with other agencies are really important. PLANETS, a European wide project on digital preservation, funded to the extent of Euros 15 million over five years by the European Commission, is one example of the international collaboration which is taking place in this area.

The next session – entitled *New Trends in Scholarly Communication*, was introduced by **Dong Youl Jeong** from Ehwa W. University in Korea. It was followed by a representative championing the cause of secondary publishing information. With the title ‘Secondary Publishing in Changing Times’ **Helle Lauridsen** from ProQuest described some of the challenges facing this intermediary group. Secondary

services, though they are built on the indexes and abstracts culled from primary publications (journals and books) in the research and professional publication sectors, are no longer the small cousins in the information industry. Recent years have seen the rise of large organisations and services in this area including World of Science (Thomson-Reuters), Scopus (Elsevier), CSA Illumina, Google Scholar, etc. Though this industry traces its online capability back some forty years, it has been in the past ten years that access to secondary services using online tools has become the norm. But still it remains largely a clone on the primary printed publication. In many cases the original fulltext publication, to which the indexes and abstracts refer, are available only as PDF documents and not HTML/XML.

There are still gaps in the secondary service provision. According to a ProQuest study, 74% of users want access to data in figures and tables. 70% want a single access point to a portal of all linked presentations (including conference reports, theses, etc.). But this is not easily achievable whilst the original publication is not accessible or technically not searchable. The key point made by Ms. Lauridsen was that researchers want access to data, to enable local computation of other results, and not just rely on images of the reported results by someone else. Even attempts by Google, through Google Images, to meet this user need is not ideal. The BioText Search Engine, feeding off PubMed Central (PMC) data, whilst meeting some of the needs, is also constrained by the sample of papers included in the source database.

CSA Illustra is an attempt to drill down into this growing area of providing supplemental information. Two million figures from 2,000 journals covering 18,000 articles are now included in Illustra. Using research results collected by Prof. Carol Tenopir and others, a proof of concept was undertaken involving seven universities and two research institutes in Europe and the United States. A highly computer literate group from these institutions were surveyed and questioned whether they thought information could be found without a table and figures search capability. The response was largely 'yes'. But equally it was felt that more time would be required to do so. And time is a valuable commodity. It was concluded that there were many benefits from having a table and figure search capability.

To test their market assumptions CSA invited sixty scientists, mainly post-docs, and who had considerable search experience, to provide data from 380 searches and 300,000 images. The question was whether the data could be found, and if found, could it be within a reasonable and acceptable time frame? The challenge was to come up with a procedure which allowed the data to be indexed with the least possible effort. The data would be both scanned images and OCRs. The images need to be pulled out from the source material. External PhDs were used to provide the index terms before an internal review was undertaken. Then there was the need to get the rights. There is a link with the Copyright Clearance Center to expedite rights payments for each individual publisher when such payments are required.

CSA launched its Illustra in March 2007 for the Natural Sciences. In June 2008 Illustra Technology was launched. This serves a different technology area and therefore has a different presentation format. They both provide evidence that the range of research material which researchers need access to is widening, and digitally-based schemes such as CSA's Illustra plays a great part in extending such visions.

**Choon-Shil Lee**, professor from the library and information department at Sookmyung Women's University in Korea was the next speaker and spoke about the medical information sector in Korea. This year represents the 100th year anniversary for the Korean Medical Association.

The Second World War and Japanese occupation meant that there was a disruption to the scholarly publication schedule during this period, from which it took many years to recover. However, between 1958 to 1995 there has been significant growth in journal publishing. In total there are now some 300 journals published in Korea.

In terms of medical articles specifically there are about 20,000 articles being published in 131 journals. *Ishosun* is a particularly significant Korean medical journal and which has been in existence since 1948. Many of the good articles are made available internationally within the mainstream western journals. There has therefore been a recent decrease, since about 2000, in the number of Korean journals largely because the competition from the established English-language international journals has become so intense. This has meant that the abstracts of most Korean research papers are increasingly being published in English. But in general Korean research results tend to be isolated because of the language barrier and accessibility problems.

Nevertheless, five of the Korean medical journals are indexed in Medline (whereas there are no journals in Science Citation Index).

There are no major commercial journal publishers in Korea, nor any semblance of a private sector e-journal hosting/publishing organisation, and there are very few subscription-based titles. In 1996 an umbrella learned society was established to provide support services to the embryonic medical publishing system. This is known as KAMJE – the Korean Association of Medical Journal Editors. This organisation provides workshops, supports the journal review processes, provides medical database support services, and assists with the distribution of Korean medical journal publications both home and overseas. There are now 168 member societies with six main committees. A Guidelines for Publication was published in March 2008.

Other recommendations made by the learned society support service was that all databases should be available in the English language. Open access to Korean literature should also be made searchable, and they should conform to relevant international publishing protocols. Open access is being promoted heavily within the Korean community, in part because there are no commercial publishers with which to compete in terms of appropriate business models.

KAMJE has also been effective in creating a bibliographic database feature for Korean medical information. This includes services such as KoreaMed, KoMCI and Synapse. The KoreaMed database consisting of bibliographic data and abstracts began operations in 1997. The KoMCI citation service was inaugurated in 2001, and Synapse, a full-text and reference linking operation, began its activities in November 2007. There is an overall commitment to ensuring that Korean medical literature is accessible to the world community. To achieve this the aim is to enable all Korean medical information to be published as open access and to abide by the emerging international standards on formatting (PMC DTD and XML, DOIs, etc.).

After Prof. Lee's presentation there was a short break. The conference then split into three different tracks, with eminent speakers in each. The first of these breakout sessions was chaired by **Brian Hitson**, from OSTI at the Department of Energy in the United States. The topic on which he moderated was 'Dynamic Engines for R&D' and included a presentation by **Inseok Song**, project manager at KISTI. Other speakers in this session included **Toshiro Minami**, professor at KIIS in Japan, **Rudzhero Gilyarevskiy**, a professor from VINITI in Russia, and **Dr. Walter Warnick** who is director at OSTI, DoE in the United States.

**Inseok Song's** presentation covered the NTIS R&D Navigator. According to the speaker there is a need for a common ICT (information and communications technology) infrastructure which supports the exchange and sharing of information across all stakeholders in the national R&D domain. However, the landscape is currently bedevilled by different approaches and structures on an international basis. In Belgium there is FRIS; in Germany there are ElektrA, GEPRIS and Vascoda; in Japan it is e-RAD, and in the United States there are a number including CENDI, RaDIUS, FASTLANE and NTIS. Korea has investigated the alternative options and in 2008 adopted the NTIS approach.

NTIS is the National Technical Information Service in the US. It provides a system which builds a central information unit supporting all stakeholders' activities within the national R&D domain. Because of its federal roots, the NTIS service adopts a top-down approach. In the Korean application of the NTIS system KISTI has collaborated with two other partner institutes, 12 funding agencies, 14 R&D supporting ministries and two designated institutes for holding the R&D data. A data model as developed which identified over 340 core data elements for R&D management, and adopted a controlled vocabulary and consistent metadata standard for exchanging basic information about each R&D element.

The result is that the NTIS Portal in Korea is the main single point of entry for all R&D information searches. Besides data about the results of an R&D project, there are specialised areas of the site dealing with personnel information, experts in the field, results, facilities and equipment, etc. Statistics are now being collected on the usage being made of the NTIS portal, and in the future the aim will be to make the portal more operationally efficient to enable R&D efforts in Korea to be further improved.

**Prof. Toshiro Minami**, from the Kyushu Institute of Information Sciences (KIIS) then focused on the issue of 'the Library as a Partner in Research'. The speaker's basic message was that the next generation of librarians should become more effective partners in the research process, to assist the researcher wherever possible. A new model was proposed for the next generation of libraries.

Prof. Minami began by outlining the five 'laws' of library science which included items such as books are for individual use, that reader's time should be saved ('time is money'), and that according to S.R. Ranganathan, the library is a growing organism. But the environment facing the librarian is changing. The Google Generation is emerging with its different engagement with the library; advanced digital technology is producing new services, and content is changing, becoming more hybrid in its nature. To keep track with these external developments the library services have changed, with (in the case of Kyushu University library) computerised OPACs being launched in 1981, and web-based information services in 1997. More types of material is being digitised, from rare books to personal correspondence. The library home pages are attempting to provide access to all this new digitised material.

In the speaker's view there will be a change with the 'information concierge' and 'research concierge' emerging. The library will become a partner in the research process. It will have to ensure that an extensive collection of required media will become available and ubiquitous to the clientele. Greater education of the librarians to become almost research assistants will be needed, and a personalised, customised approach to solving information queries will have to be undertaken. There will be a premium on providing access to raw data, and the application of RFID will be necessary. Intelligent bookshelves being experimented with at Kyushu University were described, as was the intelligent 'browsing table'. Data is collected at all points in the use of information to enable profiles of research interests to be developed. All this will lead, in Prof. Minami's view, to the ultimate creation of a virtual bookshelf within the library of the future.

The All-Russian Institute for Scientific and Technical Information (VINITI) then gave a presentation on the topic 'Integration of Knowledge Organisation Systems in VINITI'. The presentation was made by **Rudzhero Gilyarevskiy**. He made the point that Knowledge Organisation Systems include a whole panoply of classification schemes, ranging from subject headings, authority files right through to emerging semantic ontologies. Though there are differences in approach each KOS stamps its own consistency on content which is critical for digital libraries.

VINITI is Russia's largest bibliographic abstracting centre, producing approximately one million abstracts a year from over 5,000 scientific journals. It has developed its own classification system, the Rubricator, for organising abstracts within VINITI's *Referativniy Zhurnal*, and has adopted the Universal Decimal Classification system (UDC) for individual abstract identification. Software has been

developed which links the various knowledge organisation systems to VINITI's Rubricator. As a result the 15 or so Rubricators that have been developed have been judged to be highly effective in enabling users reach their required information artefacts.

The semantic web, which includes the provision of information within the content which enables not only human understanding but also machine interpretation, is also being undertaken at VINITI, particularly in the area of mathematics. It was claimed that Resource Description Framework (RDF) may be inappropriate for developing highly sophisticated ontologies in some of the very specialised research areas, but Mr. Gilyarevskiy pointed out that this was not the case with the adoption of OWL (Ontology Web Language). OWL can interpret from different ontologies.

As the speaker indicated, services such as the Rubricator are extremely valuable for work in scientific and technical information. Rubricators are as necessary as the handbook, allowing a stream of queries to be synthesised. It is the switching mechanism between different classification schemes.

Anticipating an event which will become part of the ICSTI proceedings in the 2008 Public Conference was a presentation given by **Dr. Walter Warnick** from the United States Department of Energy. The title of his presentation was 'WorldWideScience.org – Accelerating Global Scientific Discovery'.

As background, the project WorldWideScience.org had its origins in the successful 'Science.gov' project which was developed by a consortium of federal libraries and information centres attached to US government agencies mainly around the Washington DC area. This is the so-called CENDIS consortium. The Science.gov project enabled end users to search through scientific and technical information created mainly by US federal research funding, and to do this without any financial or licensing barriers put in place.

However, Dr. Warnick described a meeting which was held in the board room of the British Library in January 2007 at which Dr. Raymond Orbach, Under Secretary for Science at the Department of Energy, and Lynne Brindley, chief executive of the British Library, signed a joint Statement of Intent to become partners in the development of a searchable global science gateway. WorldWideScience.org (WWS.org) was launched in June 2007 and involves participation from a large number of countries. The aim was to enable easy access to publicly funded research results which affect scientific and technological progress and innovation. The aims and scope are, therefore, similar to those set by the CENDIS group and which resulted in Science.gov. The main difference being that WWS.org would be global in coverage and content.

During the past eighteen months WWS.org has grown from 12 databases accessible through the system from ten countries, to 32 databases from 44 countries. Every continent is involved in partnering with WWS.org, and 200 million pages have been brought together. Dr. Warnick listed all the member countries and organisations which are participating in the global service in his presentation. He also described the governance of the project, with Richard Boulderstone (The British Library) as chair, and Ms. Pam Bjornson, from CISTI in Canada, being elected as Deputy Chair. The Treasurer is Tae-Sul Seo, from KISTI in Korea.

In terms of technology, many of the major current search engines rely on intermittent crawler technology to collect and index material which is then accessible through their respective services. In the case of WWS.org a different approach was adopted – a federated search engine was developed which went below the surface web and collected information from the much larger deep web. The search software was developed by Deep Web Technologies in Arizona, USA. As such there will be material available within WWS.org which may not be readily accessible through Google, Yahoo and the other specialised search engines.

The speaker also commented on some future enhancements which are being explored. Results clustering is one of these as well as the provision of personalised alerting services. In view of the international nature of WWS.org it is also proposed to investigate the potential to include automated translation capabilities.

The second track which occurred simultaneously with the above track was about 'Data Management and Integration' and was chaired by **Tony Llewellyn**, consultant and advisor to ICSTI. The speakers in this session included **Kum Won Cho**, a team leader at KISTI in Korea, **Dr. Lee-Ann Coleman**, head of STM at the British Library, **Alex Wade**, senior research program manager at Microsoft Research, **Peter Stoehr**, head of the Literature Service at EBI, Cambridge, UK and **Hyung-Seon Park**, project manager at KISTI.

KISTI's e-Science Division produced a joint paper which was presented by a member of the team, **Kum Won Cho**. The presentation was entitled 'Scientific Data Management in a Global e-Science Environment'. It raised the issue of 'what is e-Science'? The answer given by the speaker was that it involved a high-end cyber research environment which increases the productivity of research by integrating high performance computing facilities, large scale data, human resources, etc., through the provision of leading edge ICT and high performance networks. The speaker then proceeded to put this in the context of KISTI's e-Science activities.

KISTI has a branded service which is Korea's global science gateway – K e-Science. It is a multidisciplinary global collaboration which supports virtual communities in physics, life sciences, engineering and geosciences. The infrastructure is international but it is supported by KISTI's fourth supercomputer. In terms of applications the speaker described the details of one such service – e-AIRS. e-AIRS is focused on the fluid dynamics research environment (e-AIRS is Anytime, AnyPlace, Any Fluid Research). It includes remote experimental services (wind tunnel), video collaboration, grid computing, etc. A detailed expose of the technical architecture for e-AIRS was given.

Other K e-Science applications which were described included the CFD Solver service, the POSE Parametric Study Service, WISDOM (Wide In Silico Docking on Malaria), and a remote conferencing tool. The speaker also referred to the work which KISTI was able to do in monitoring the outbreak of Avian Flu in 2008.

Included in this session was a presentation by **Dr. Lee-Ann Coleman**, Manager of the recently created STM department at the British Library. She described a particular new project with which the British Library and Microsoft are cooperating on in bringing to market. It is referred to as the Research Information Centre (RIC) and is planned to become a holistic desktop information system which addresses the extremes of information needs of scholars and scientists. An early attempt was tested at the University of Leeds, and from this an IT-sustainable technological solution, built around Microsoft's SharePoint system, has emerged. The whole life cycle of the research process will ultimately be included in a system with a single point of access.

The basic premise is that many new media forms of communication are not included within the traditional remit of the library. RIC has therefore been developed to address the various emerging 'pain points' in the scientific research process. It has to include an adaptable framework for end users. The biomedical community is the first of the scientific communities which is being targeted with RIC.

The main functions which have been included within the RIC framework include resource discovery (of not only published literature but also datasets and the Internet itself); alerting (providing RSS feeds on newly published literature but also providing alerts about new sources of funding); collaboration (social networking in the drafting of proposals, messaging and document sharing); personalisation; facilitating

experimentation (in providing access to tools and data manipulation); and dissemination (which aids document production, uploading manuscripts to the publisher and deposition in a repository).

So far the experimentation phase is the least well-developed of the current functionalities within the system, however, these and other functions are due to be added as and when market testing shows them to be appropriate. In order to ensure that the RIC service is what the market wants some 18 research groups have been recruited (14 in the UK, the others throughout the world) as part of a closed beta evaluation. Once the feedback has been received from these groups and adaptations made the service will move on to become an open beta test.

Following on from the British Library interpretation of RIC, **Alex Wade** from Microsoft Research (MSR/External Research) then commented on the project from the Microsoft perspective. The aim of Microsoft Research is to transform the process of scholarly communication and make it more efficient. He described how Microsoft views the research life cycle and the information life cycle. For example, he pointed out that conference management tool also needs to be included in any effective information system.

Alex Wade focused on what he defined as the four main phases of RIC. These were the same as those outlined above but include:

#### Phase one: Idea, discovery and design

- Integrated search
- Automatic paper download
- Search of the archive
- Personal and project reference lists
- Personal and project paper libraries
- Paper recommendation
- Continual query service through alerts
- Bookmarking, tagging and annotation tools

#### Phase two: Obtain funding

- Proposal templates
- Examples of successful proposals
- Search for funding opportunities
- Proposal review

#### Phase three: Experimentation and analysis

- Search for tools
- Rating tools
- Annotation tools
- Research output libraries
- Common tools for STM

#### Phase four: Dissemination and publishing

- Version management
- Review workflow
- Bibliographic management
- Support for wikis and blogs

- Share updates via RSS feeds
- Provide copyright compliance tools (CC)
- E-journal management tools

**Peter Stoehr**, from the European Bioinformatics Institute (EBI) near Cambridge, UK, provided an update on the latest developments in bioinformatic information and how this accentuates the role of original raw data as a primary information source rather than just relying on published articles and books. In passing he mentioned that the role of EBI is to develop structured databases of biological data representing nucleic acid and protein sequences, protein families, macromolecular structures, microarrays and others. Despite these efforts, Peter Stoehr stressed that most biological knowledge remains in the text of scientific literature which underpins these biological databases.

It is apparent that bioinformatics is nevertheless well served with digital information services. They range from EB-eye/Entrez browsers through CiteXplore, which is an interface to a number of significant biological databases launched in December 2006. It includes links to fulltext through EMBL-Bank. It includes pictorial databases such as PDSum, providing an overview of every macromolecular structure deposited in the Protein Data Bank.

Many of these services have well-established links between the published literature and databases. But in these instances access is often through accession numbers rather than accessing the data itself. In many cases scientific articles in the bioinformatics area will not get published unless the raw data supporting the research is deposited and made accessible from a recognised dataset (which includes the accession number).

Peter Stoehr also discussed the text mining services which are seen as being important to reveal some of the concepts currently hidden within and between datasets. He referred to tools such as Whatizet, which undertakes a translation into a database structure, and EBI Med which searches for words in other contexts, plus several others. Whatizet is graphically appealing in that it uses colours and hyperlinks to effect switching between different databases. He also commented on the available search engines which are appropriate to the biomedical world. These include EB-eye, an Entrez browser, which tells the researcher in which database a particular term is located.

A description was also given of PubMed Central, the free archive of biomedical articles maintained by the NCBI at the National Institutes of Health in Bethesda, USA. Some 50,000 articles are available, in XML, to read and download for free. Some of the older articles are scanned images and some are also in HTML/PDF. EBI is currently involved in a project, together with the British Library and the University of Manchester, to replicate PubMedCentral in the UK. This project, funded by the main research agencies and the Wellcome Trust, is aimed at providing local support for the creation of such a free database as well as enabling the UK funding situation to be taken into account in the structure and governance of the service. Other countries, including Canada and Japan are positioning themselves to set up similar national mirror sites of PMC, and an European version is also being investigated.

The final speaker of this track was **Dr. Hyung-Seon Park**, from the Korean Biodiversity Information Facility (KBIF). His presentation was on 'The biological diversity information flow in KBIF – the activity and the role to GBIF'.

The Korean Biodiversity Information Facility is part of an international collaboration involving 48 countries and almost 40 organisations with the aim of making the world's biodiversity data freely available via the Internet. This programme is the Global Diversity Information Facility (GDIF) with its secretariat based in Denmark with 17 full-time staff. The sort of information collected within the GDIF includes taxonomical data, specimen data, characteristics of molecular genetics, ecological and taxonomical systems, etc. It is a vast, complex and significant area. Yet there is still much work to be done.

GBIF started life as part of an OECD Working Group on Biological Informatics in 1996, and which became operational in March 1999. A central format for data exchange was developed, known as Darwin Core 2, which acts as a link between the challenges facing biodiversity which is in the developing world, whereas the data which is available to monitor this resides mainly in the developed world. The GBIF brings together collaboration from museums, universities and observers throughout the world. GBIF is ostensibly a global partnership. But still only 4.5% of available records are shared – ‘95.5% are sleeping records’ according to the speaker. Mirror sites of the data records are available on three continents – USA, Denmark and at KISTI in Korea.

KISTI, with its KBIF unit, undertakes the whole range of local support services for GBIF including training, information dissemination, portal development and running the local data repository.

The third and final track for the first day’s proceedings was a special session on licensing issues in electronic information. This was chaired by **Jeong Hwan Kim** of KISTI. It included a detailed presentation by one of the world’s experts on SUSHI (Standardised Usage Statistics Harvesting Initiative) – by **Oliver Pesch**, chief strategist at EBSCO, USA. An additional presentation in this session was given on the KESLI – Science Direct proposal for the year 2009 by a representative from Elsevier Korea.

**Oliver Pesch** from EBSCO Information Services provided ‘A beginner’s guide to NISO’s Standardised Usage Statistics Harvesting Initiative’ (SUSHI). He explained that this initiative, SUSHI, was important because librarians have a need to collect statistics on usage – their funding bodies and government agencies demand it of them. The usage data informs on renewal decisions, cancellation decisions, cost per use, and generally provides for the management of e-resources within the institution. He described the historical development which led to the NISO protocol on SUSHI – starting with COUNTER being formed in 2001, to the emergence of ERM and the ERMI committee in early 2002, culminating in the formation of the SUSHI committee in mid-2005. He explained that COUNTER provides the mechanism to ensure credibility and consistency of the collected usage data, ERM provides for the consolidation and meaningful reporting of the data, and SUSHI enables the retrieval and processing of usage data reports.

COUNTER has a number of Codes of Practice. The first was a code of practice for Journals, launched in January 2003 and with the most recent update in January 2006. Books and Reference works were covered in a code of practice released in March 2006. Each of the codes provides data on requests, turnaways and total searches in a highly structured and comparable way. Consistency and Credibility are the bywords. It is also planned to release a version 3 of the Journals and databases code of practice in the near future which will deal with consortia, with federated searching and with robots.

SUSHI has emerged because librarians need a more efficient data exchange model. SUSHI is a web-based service model for requesting usage data, thereby replacing the user’s need to download files from the vendor’s websites. This is facilitated by the knowledgebase of the institution’s own collections – the ERM knowledgebase. An ‘import’ function enables remote data to be collected from content providers. A reports interface then is used to select the scope of the required reports. However, doing this for each content provider can be a time-consuming activity. The role of SUSHI is to allow the consolidation to take place without human intervention.

An XML schema (XSD) is behind the application of SUSHI – this is a widely supported (WWWC) and is secure. It is an open standard that interacts with other web applications for exchanging data. A Simplified Object Access Protocol (SOAP) is applied. This is a lightweight XML based protocol for use on the web. This enables two totally different web-based information services to be interrogated through a SOAP messaging format.

Oliver Pesch then described how a SUSHI transaction took place in practice.

The second day of the conference also started out with three tracks. The first was focused on STI Service Innovation in East Asia and was chaired by **Young Man Ko** from Sung Kyun Kwan University in Korea. The speakers in this session included **Jiachang Chen**, Deputy Director of ISTIC in China. This was followed by **Yukiko Sone**, director of JST, Japan and then by **Tae-Sul Seo**, Senior Researcher at KISTI.

The second track was about Scientific Data Management and Integration, and was chaired **Hye Rhan Chang**, from Sangmyung University in Korea. The speakers in this session included **In Soo Ko**, senior advisor at PAL in Korea, **Sang-Ho Lee**, project manager at KISTI and **Jan Brase**, project manager at TIB, Hannover in Germany.

The third track was a special session on the Evolution of Information Services in Portals, and was chaired by **Jeong Hwan Kim**, from KISTI. Speakers in this session discussed the proposed KESLI – World of Science (WoS) proposal for 2009 as given by a representative from Thomson Reuters in Korea. The product manager for Google Scholar in Korea (**Haimin Lee**) also gave a presentation on this particular Google service offering, and **Beom-Seok Seo**, division head of NHN in Korea completed the proceedings.

The Deputy Director of ISTIC in China, **Chen Jiachang**, led off the first track with a description of ‘The Scientific and technical information Service and support system in China’. In China the government has put forward a national policy of integrating IT applications with industrialisation. In this policy the development of an information society plays a crucial role. Innovation is also highlighted within China’s long-term development strategy.

The first phase of the information society programme was launched in 1992 and involved focusing on computer applications and technology. Phase two followed in 1993–1997 and included the identification of three so-called ‘Gold Projects’. Phase three (1997–2000) established a long-term vision of the information society. The fourth phase ran from 2000 to currently and involves projects which will support the creation of an information society which meshes in with China’s Tenth National Five-Year Plan.

Some of the projects which are included with the State Council of China’s e-Government development include two networks (a government intranet and government internet), a portal which provides government information, four key databases (on issues such as population, corporations, natural resources and macroeconomics). 12 Golden Projects are now being developed which include collections of information on taxation, finance, government information sharing, public security, labour, agriculture, water resources, etc.

The e-Government portal opened for business in January 2006 and by the end of 2007 more than 23,700 government domains were included and 12,000 web sites were accessible. Within e-Education, there are campus networks at 90% of colleges and universities, and even 60% of middle and primary schools had site networks. Distance Learning is also a significant feature of the e-Learning programme with over 4.2 million Chinese currently enrolled in distance learning for higher education. The government is also encouraging the adoption of e-Commerce models within the industrial and commercial sectors.

Within Science and Technical information there are a variety of support systems. There are a number of main library systems, including the National Library, the National Science Library and the Institute of Scientific and Technical Information (ISTIC). There are also a number of digital libraries funded by various government ministries.

A particular STI project is the science data platform for earth systems which has been established. It supports some 310 national research projects with 3,500 customised information services. There are

similar resource sharing platforms for medicine and natural science and technology. A large-scale scientific instrument sharing platform has also been established covering 14,000 scientific instrumentation sites and facilities.

Underpinning these various projects there is a centralised consulting service to help with policy setting. Since 1986, when a decision was reached by the National SoftScience Symposium to promote scientific and democratic decision making, some 2,500 think tanks and research institutions have blossomed. 5% of these are in private hands and includes the Horizon Group and Greatwall Strategy Consultants.

China is developing rapidly in its online information usage, with approximately 200 million online users in 2007. A good entrance to what online information is being generated and available within China is through the China info web site.

ISTIC itself consists of a number of functional divisions. These include:

- The Information Resource Centre
- Strategic Research Centre
- Resource Sharing and Promotion Centre
- Information Technology Centre
- Research Centre for information methods (which includes scientometrics)
- Postgraduate Programme Centre

ISTIC is the main Chinese focal point for establishing and maintaining relationships and partnerships with agencies such as UNESCO, IFAP (Information for All) and ICSTI. Besides the various library support services which it coordinates it also offers a document delivery service. It is active in the application of digital object identifiers (DOI) in China.

From Japan, the JST representative was **Yukiko Some**, who gave a presentation which focused on Japan's national plan, the Third such plan, on 'Creating Innovation'. The national plan involves the commitment of \$16.2 million to R&D in the areas of innovation, of which 80% is due to come from the private sector. Currently Japan is responsible for creating 10% of the world's scientific/technical and medical literature. It is fourth in the world in terms of citations, even though many papers have a low citation count. In some subject areas it is stronger than in others – it has above its overall average in areas such as materials science, physics, pharmacology and chemistry. It is comparatively weakest in ecology, environment, mathematics and geosciences.

The main source of published articles emanating from Japan are commercial publishers worldwide which are responsible for 65% of the output. 7% of these are from domestic Japanese commercial publishers. 32% are published by learned societies and 7% come from university presses. There is an imbalance in the research information system – the domestic publishing scene is too small and highly journal-dependent. Japanese journals in the Japanese language has a limited global market.

There is some support available through J-STAGE, however. J-STAGE is a platform supported by JST to encourage academic societies to publish e-journals for international circulation. It supports publishing in both English and Japanese. It currently includes about 500 journals. In terms of access to the J-STAGE online service, 40% comes from Japanese researchers and 14% from China. The US is the next most significant online user with 10% of accesses. In addition SPARC operates in Japan – 45 journals are currently included within this support service. There is also an archiving service available to Japanese publishers, Journal@rchive, which is also hosted by JST and contains the backruns of almost 200 titles one of which dates back to 1877 (Physical Society of Japan).

Japan is also involved in the open access movement, with the 3rd Science and Technology Basic Plan recommending that free open access to the results of publicly funded research be offered. According

to the speaker there are 360 journals included within the J-STAGE platform which are available for free. Japan also has invested in the (Green) institutional repository programme, with 81 universities and institutes having their own established repositories.

Yahoo is the dominant search engine in Japan (Yahoo JAPAN), with almost 60% of usage, and Google has a further 32%. However, there have been a number of recent initiatives which provide a greater link between Google and the Japanese academy, including GoogleScholar and Google targeting J-STAGE, and with Keio University being a partner within Google Book Search.

Another Asian country, this time the host nation, then described the Innovation of S&T Information Services in Korea. This was undertaken by **Tae-Sul Seo**, Senior Researcher at the Korea Institute of Science and Technology Information. The online services have had a variety of structures over the years, with KISTI emerging as the named service in 2001. The KISTI portal became operational in 2005. It includes access to a variety of databases provided under the auspices of the National Digital Science Links (NDSL). Features of the portal include convergence between paper and patents; it operates as a reference link service, and is international in its orientation. The KISTI portal includes access to WWS.org (see earlier). It also includes a number of Web 2.0 services such as 'my NDSL', tag clouds, community facilities, blogs and evaluation.

Following on from this expose of the overall Korean online efforts, **Prof. In Soo Ko**, of PAL, POSTECH, described 'Scientific Data Management at Pohang Light Source'. Pohang Light Source (PLS) is the largest research facility in Korea covering a broad range of research areas, ranging from basic science to biomedical to engineering applications. PLS has become a flagship for the development of Korean science and technology. The PLS facility includes two major accelerators. The information system developed at PLS, whilst not large in terms of throughput, is unique in that there is a need to capture the information in real time and used as an input to various feedback systems.

Following on from the approach adopted for highly technical/engineering information, the next speaker focused on the role of factual databases in Korea. This was also presented by another representative from KISTI, in this case **Sang-Ho Lee**. He described a number of databases available to Korean databases including a Chemical Database, a Plasma Property Database, an Inorganic Crystal Structure Database. In addition there is Digital Korean, which covers medicine, biomechanics, human modelling, industrial safety, etc. Then there is Visible Korean for medical education, a Korean Traditional Knowledge portal and a gateway website. The speaker went into detail about the content of each of these factual databases.

**Jan Brase**, from TIB Hannover, gave a presentation on the application of digital object identifiers (DOIs) to raw data. He described the sequence of events which went from the production of raw data within an experiment, through compilation of information, to the publication of the analysis, to developing a knowledge base. This sequence is only structured and formalised in the middle ranges of the spectrum. There is limited focus on providing structure, accessibility and preservation of the raw data. Yet without such a structure there is a risk of unnecessary duplication of research results, and the social waste which this can lead to. There is a need to create validation of research claims through the raw data. There is a great deal of data out there but we do not know what is there and what is significant.

According to Jan Brase what we need is better, more sophisticated and professional data centres. These should ensure global access to the research datasets. Each of these datasets should be identifiable through a unique DOI. This would enable them to be citable in the literature, rather than someone's interpretation of them in published articles which is the current situation. We need citable primary data. If the raw data becomes citable then the citation potential for the research paper itself can become greater.

In Germany data is stored for ten years. This enables a researcher to check back to see if the research has already been conducted and what the results were. This can and does stimulate new research ideas – to fill in gaps which the existing raw data sets do not touch. It also allows longitudinal studies to be undertaken and the historical data sets to be identifiable and easily accessible through DOIs.

Some of the existing datasets within the German system include research results stemming from the Deutsche Forschungs Gemeinschaft (DFG) research funds. TIB is also involved in applying DOI principles to data stemming from world climatic data centres. There are many partners involved in this. Data which is sent to TIB for DOI application must include appropriate metadata which describes what the raw data consists of. TIB then sends it for DOI application and access to it is then resolved through the Handle system. So far this system has been community funded and now contains 550,000 objects.

TIB itself does not store the datasets – they are stored and preserved by the various partner organisations. As far as the large datasets are concerned TIB does not have the expertise to manage these datasets – again, reliable or ‘trusted partners’ are identified and they adopt storage and preservation procedures which are in line with enabling access to the data to be achieved. DOIs can also be initiated in the work process itself.

DOIs might be seen as part of the ‘open access’ movement in that data is held to be free, and DOIs of the data will aid in its exposure, accessibility and use. However, DOI itself is not without a cost. It is a commercial registration system governed by the International DOI Federation (IDF). There are licences involved, people to check metadata, etc. Currently TIB gets its money to support DOI identification of datasets from three main sources:

- DOIs paid for by the publishing agencies.
- TIB itself as part of its national mission. TIB is funded by the German Ministry of Research.
- Other funding agencies in Germany.

Datasets will be presented the same way as literature – abstracts of data (metadata), DOIs, etc. – and are accessed through a portal page at TIB. The DOIs applied by TIB will be included in the CrossRef literature referencing system created and maintained by the main journal publishers. TIB would like to see more library involvement – so far subject based data set agreements have been reached with ETH (Zurich) and INIST (Nancy). ICSTI involvement is also being evaluated.

But this was not the end. After the mid morning break the delegates joined together for a final plenary presentation given by the president of OCLC in the USA.

**Jay Jordan**, president and CEO of OCLC, described the activities in which OCLC, the leading international library support services group, is currently involved. OCLC has undertaken many studies in the past and Mr. Jordan summarised some of the results from these. They relate to libraries in the world of scholarly publishing as well as libraries in the community in general. The big issue which has emerged is the new elephant in the room – Google – and what this means to libraries. In some respects libraries are absent from the significant debates which Google has spawned. He pointed out that 84% of users rely on search engines to begin an information search whereas only 2% use the library web site. 90% of users are apparently satisfied by the existing search engines, whereas the general conception seems to be that libraries are a place for books. Jay Jordan tried to identify what OUR (meaning OCLC) space was. This was the culmination from four major strategic reports, one of which is due to be published in a few weeks time.

OCLC operates worldwide. 60,500 institutions in 112 countries are its main clientele. 25% of its revenues come from overseas. In addition, OCLC has 26 offices of its own in ten countries.

From the results of research conducted by a number of agencies it appears that 25% of the general public and 50% of college students, participate in social sites. The WorldCat has a link into FaceBook, and the API for WorldCat is generally available for social networking/collaboration purposes. A unique URL can be created for each item which enables items to be accessed from individual libraries. A Virtual International Authority File is also being created with the help of international partners such as CNRS. Mr. Jordan showed a 'map' which illustrated the interconnectedness between the 200 most successful web sites. In order to ensure the relevance of the international cooperative OCLC undertakes strategic planning exercises, and the key messages from the 2000 Strategic Plan were to extend the cooperative, to introduce new services, transform its central cataloguing system (WorldCat), build a new technological platform and weave libraries into the web. Mr. Jordan then listed the various achievements which have been made along the lines laid out in the Strategic Plan.

In addition OCLC has worked in other areas as well. It has established its own Digital Archive service for libraries, enabling them to manage and preserve their digital collections. OCLC also acquired NetLibrary several years ago which enabled OCLC to exploit the technology to create and sell e-Books. Other specific new services include creating a WorldCat registry, establishing 'QuestionPoint' jointly with the Library of Congress to provide answers to users' questions. However, OCLC's main focus remains to catalogue the world's records, and by this year has 100 million records catalogued. Mr. Jordan listed the key libraries which have contributed to the WorldCat catalogue – the Godfrey Memorial Library (with 135,800 records) is the largest – and he also listed the various national centres which were in the processing of adopting WorldCat. This is part due to the multilingual nature of WorldCat, with 45 million records in English, nearly 10 million in German and 5 million in French.

As a result WorldCat currently receives 30 million users a year, hitting 84 million records. These will be enhanced as developers use WorldCat API to create applications that will drive additional users back to OCLC services. This is particularly the case as WorldCat Local Pilot which enables interoperability with local systems in offering resource sharing and resolution to local holdings of fulltext. One example quoted by Mr. Jordan was that the University of Washington's use of the WorldCat Local resulted in a 100% increase in inter-library loans during 2006/2007.

OCLC Research has focused on a number of projects in the areas of content management, collection and user analysis, interoperability, knowledge organisation and systems and service architecture. For example, building communities and convening groups of experts is being coordinated through the RLG Programmes. Book digitisation projects are being undertaken with Google, Gutenberg, the Open Content Alliance and the Universal Digital Library. He commented briefly on the 'commercial unsustainability' of the recently deceased Microsoft project. However, OCLC and Google have an agreement to exchange data which enables those books which are digitised under the Google book digitisation programme to be included within WorldCat, thereby enhancing their use. This is a reflection of a fundamental digitisation policy advocated by the American Library Association which says that no one organisation will have all the answers and that large-scale digitisation requires national and international collaboration. In this respect Mr. Jordan pointed to the work being done by the Blue Ribbon Task Force on Sustainable Digital Preservation and Access which aims to push back the frontiers of digitisation on a coordinated basis relying on the experience of noted experts who have been appointed to the committee.

Mr. Jordan completed his presentation by referring to the quotation from Charles Darwin which, paraphrased, points out that it is not the strongest nor the most intelligent of the species which survives, but rather the one that is most adaptable to change.

### **ICSTI Collaborative Project**

Another significant development was then announced to the attendees and an invited audience also from the media. It was the formal launch of the **WorldWideScience Alliance** (WWS.org) by means of a signing ceremony attended by some of the key players and funders of the service. The WorldWideScience Alliance is a group of international organisations which have collaborated on the creation of a database of mainly scientific, technical and medical information which will be freely accessible throughout the world. It is strongly embedded in the ethos which received such strong support during the ICSTI Conference that there should be international collaboration in order to promote innovation of STI. In this instance the origins of WWS.org lie in work which federal libraries in the United States undertook several years ago to enable a single point of access to their combined information resources. Some 15 different federal agencies were involved, and Science.gov was the single access point to their individual resources. The constituent databases were searched through a federated search system developed for Science.gov by Deep Web Technologies.

WWS.org builds on this framework and public funded and national organisations were approached each with a strong science base of information and an aim to disseminate their information as widely as possible. The lead organisation behind WWS.org was the Office of Scientific and Technical Information (OSTI), a division of the US Department of Energy. In January 2007 an agreement was signed by the director of Science and the chief executive of the British Library to promote the WWS.org concept, and ICSTI as a body endorsed the service which has since become part of the ICSTI programme. The formal launch ceremony including a ceremonial signing of the relevant agreements to create and sustain WWS.org was the concluding session of the ICSTI 2008 Public Conference.

**Dr. Walter Warnick** from OSTI/Department of Energy presided over the signing ceremony. The ICSTI representative also added his pleasure at hosting this formal launch ceremony. Representing the US Department of Energy at the signing ceremony was **Dr. Jeffrey Salmon**, Associate Under Secretary for Science at the United States Department of Energy. He described the historical aspects to the creation of WWS.org and gave strong support for its open access mission. Support for the concept was also given by **Richard Boulderstone** for the British Library who besides reflecting on the BL's early support for the programme also agreed to become the first chairperson of the WorldWideScience Alliance. Representatives from some of the other national information centres which have provided information accessible through WWS.org were also in attendance at the signing ceremony and were introduced in turn by **Herbert Gruttemeier**, ICSTI president before opening the session to photographers and the media.

The ICSTI 2008 Public Conference was an undoubted success. It puts ICSTI back on the map as an organisation which has a strong programme to create global collaboration in the promotion of scientific information creation and sharing. ICSTI has incorporated this as part of its new strategy for the future, a strategic document which was overseen in its development by **Tony Llewellyn**, advisor to ICSTI.

With such a successful conference behind it, with a longterm strategy with which to take the organisation forward, and a highly visible project in the form of the WorldWide Alliance to achieve global visibility and credibility, the auguries are good for ICSTI in the future. This framework was then built on during the various ICSTI Committee meetings which were held at the Codex Conference Centre as new initiatives and plans were considered for ICSTI for the future.