International Training Workshop for Developing Countries on Scientific Data Management

Training Initiative Proposal

February 18, 2011

CODATA-China

Chinese Academy of Sciences

1 Objectives

The workshop offers state-of-the-art scientific data management methods and information technologies for research community members coming from developing countries. Also we hope to make it become a composition of ADMIRE(Advanced Data Methods and Information Technologies for Research and Education) initiative according to the CODATA's Strategic plan 2006-2012, thereby strengthen the communication and cooperation between the developing countries and promote better known the work and mission of CODATA.

2 Brief background

Scientific data has been the core position of the revolution of research methods and e-science, CODATA will have been making significant technological and institutional impact in the application of data, many developed countries enhanced the research data management and training recently years, particularly strengthen the global education and training for data management in the future, which all consider CODATA will play important role. CODATA-China as an important member of CODATA, communication and cooperation on scientific data within CODATA and members have been promoted year by year. For example, Chinese delegation participate in CODATA 2010 conference with more than 40 scientists and graduate students,

CODATA China and CODATA US annual roundtable meeting hosted by each other has successful been held since 2005. And also CAS has been developing scientific database towards long term data accessibility and application. CAS obtained plentiful experiences and applied technologies in scientific data infrastructure and data resource management. there are some successful stories about scientific database building and data application among institutes of CAS especially in the bio-science, geo-science and materials science. Here by the opportunity of exchange experience and promotion communication on data management come into being with the CAS and CODATA channel.

3　Training organization

This Training workshop will be organized by CODATA, Chinese Academy of Sciences and the Chinese National Committee for CODATA jointly. The Chinese National Committee for CODATA will host the workshop with Chinese Academy of Sciences together. A steering committee will be established by CODATA, CODATA-China and CAS. And a work group will be established to prepare for the workshop after the proposal is approved. The 20 day's workshop holding about 20 trainees will be held in China, 2012. The rough work plan for the workshop is illustrated in section 4.

4　Work plan

Time	Work
Feb.2011- Mar.2011	✓ Discuss and refine the proposal and make it approved by CODATA
Apr.2011 – May.2011	✓ Establish the steering committee ✓ Establish the working group
Jun 2011-December 2011	✓ Select training topics, trainers and trainees from the list recommended by CODATA national members ✓ Make the training materials ✓ Other detailed work for the workshop
May 2012	✓ hold the workshop in China

5 Suggested training themes[1]

1) Technical and institutional issues in Data management

   ✓ Open access to research data

   ✓ Data sharing

   ✓ Data publishing

   ✓ Data citation

   ✓ Data curation, long-term data preservation

2) Scientific data infrastructure development and service

   ✓ Distributed Metadata management

   ✓ Techniques for replica management

   ✓ Federation of diverse databases, data mediation services

   ✓ Virtualization of data resources

   ✓ Support for streaming data

   ✓ Database performance analysis, algorithms, and data structure on modern hardware

   ✓ Linked scientific data

3) Best practise of data management

   Will be decided according to the recommending list.

6 Financial requirement

The travel support for all participants will be solicited from CODATA National Members and partner groups, other charges including the local event expenses could be supported by CAS.

---

[1] Content in this section is based on the existing data management education and training initiatives