

1st IMBER Data Management Committee (DMC) meeting

The first meeting of the DMC was held in Victoria, Canada on 10/11 June immediately before the IMBER SSC meeting. The challenge for the DMC is to design a data management strategy for IMBER that spans the wide range of multi-disciplinary data that will be collected, while countering the apathy that most researchers display towards data management. A range of incentives was discussed. A few journals already refuse to publish a paper until the data are available with the article or in a recognised database. Funding agencies may refuse to accept project reports (and future funding) until data are deposited in a recognised database. These “sticks” are effective, but the DMC wishes also to promote “carrots”, including access to experts (National or Specialist Data Centres) to help with calibration and validation, immediate backup to safeguard data and facilitating access to other data. One carrot which SCOR is investigating is the possibility of tagging a data set with a Digital Object Identifier (DOI), which would allow the data set to be referenced and credit given to the originator, just as is now done for published papers.

However, the DMC believes that it will take more than these measures to change attitudes. Funding must be allocated to data management, and researchers educated to its value. A leading data centre has found that it requires 6-12 person-months of staff effort to manage the data from a major research cruise. This includes creation of metadata; working with data-

collecting scientists to obtain, clean and calibrate their data; merging data (for example, bottle data sampled and analysed for different properties by different scientists); and putting data into a database. The solution proposed by the DMC is that a cruise (or project) PI should identify or hire a person (“Data Scientist”) to carry out this task. This could be a recent graduate or young scientist, who would be given training by a Data Centre in what is needed. The DMC will create a manual covering what the Data Scientist needs to do, which will also be a guide for PIs.

The first important step is to ensure that IMBER data are recognised and documented with suitable metadata. The DMC recommended to the SSC that all cruise reports should be accompanied by a digital form-based summary report in the IOC-agreed Cruise Summary Report (CSR) format (translators from other, local formats will be developed) and that a Directory Interchange Format (DIF) record should be created for every IMBER activity that results in a data set. The IMBER website will link to the Global Change Master Directory (GCMD). The manual will provide help with CSR and DIF creation.

The meeting report will be made available as soon as possible on the IMBER website, along with guidelines for projects wishing to be endorsed by, or associated with IMBER.



Participants to the Data Management Committee meeting from background left to right: Reiner Schlitzer, Ed Urban (SCOR Executive Director), Raymond Pollard (DMC chair), Todd O'Brien, Wilco Hazeleger, Jay Cullen, Julie Hall (IMBER chair), Sophie Beauvais (Data Liaison Officer), Gwen Moncoiffé, Toru Suzuki.