Data Management in the IODP

by Bernard Miville, Emanuel Soeding, and Hans Christian Larsen

The Challenge

The Integrated Ocean Drilling Program (IODP) is challenged to develop a program-wide, state-of-the-art data management concept enabling the receipt and handling of scientific data from three different IODP drilling platforms, legacy data from the Ocean Drilling Program (ODP) and the Deep Sea Drilling Project (DSDP), and post-drilling scientific data published in the open literature. Data collected from boreholes and cores include various types of measurements, images and downhole logging data and ranges in nature from microbiology and micropaleontology to geochemistry and physical properties. The data underpin research spanning from fundamental geodynamics and solid Earth cycles over biological evolution, the deep biosphere, paleoclimatology, and hydrosphere–lithosphere interaction. A proper data management system for the future must present these highly diverse data in a transparent way through a single user-interface and provide links to related data bases (Fig. 1). Accomplishing these goals will be a major step forward not only in ocean drilling, but ultimately also in the broader field of Earth sciences. To be successful it requires strong coordination and involvement of the IODP science and drilling operators and eventually strong integration with related initiatives in the broader community.

The IODP Management International (IODP-MI) is charged with providing the overarching IODP data portal and integrating all data provided by the different platform operators and eventually by individual researchers. The IODP-MI is also taking a lead in establishing a Web-based system for scientific drilling proposals and an all-digital data bank for pre-drilling site-survey data. In all of these efforts, actual programming work is outsourced as much as possible to interested partners in the community or via commercial vendors identified through a process of competitive bidding.

Scientific Earth Drilling Information Service—SEDIS

The key to providing easy and seamless top-down data community access from a web of distributed operator databases is defining program-wide standards in data and metadata collection. The Scientific Earth Drilling Information Service (SEDIS) data management system that the IODP-MI is pursuing builds on such metadata and will provide a unique entry point to locate all IODP and legacy data. The ISO 19115 metadata standard has been chosen in anticipation of its widespread future use.

SEDIS will be developed in close cooperation with the IODP operators and other Earth science initiatives, and will
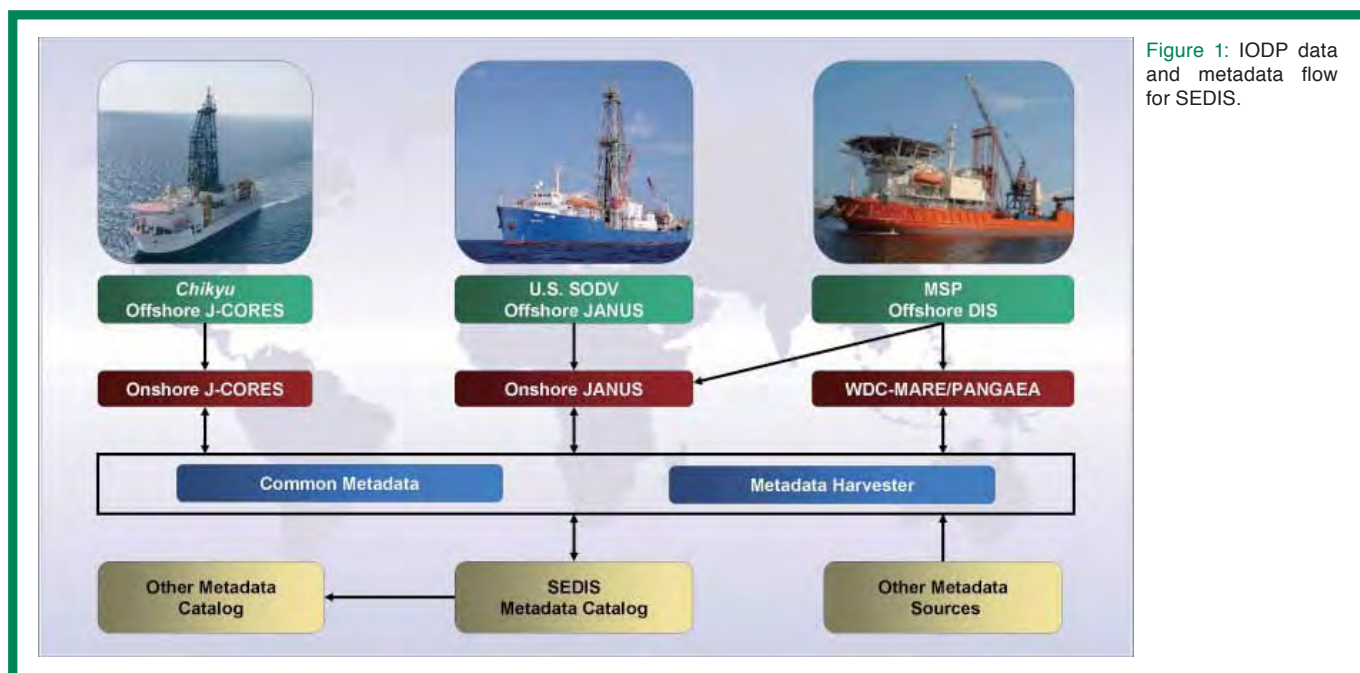

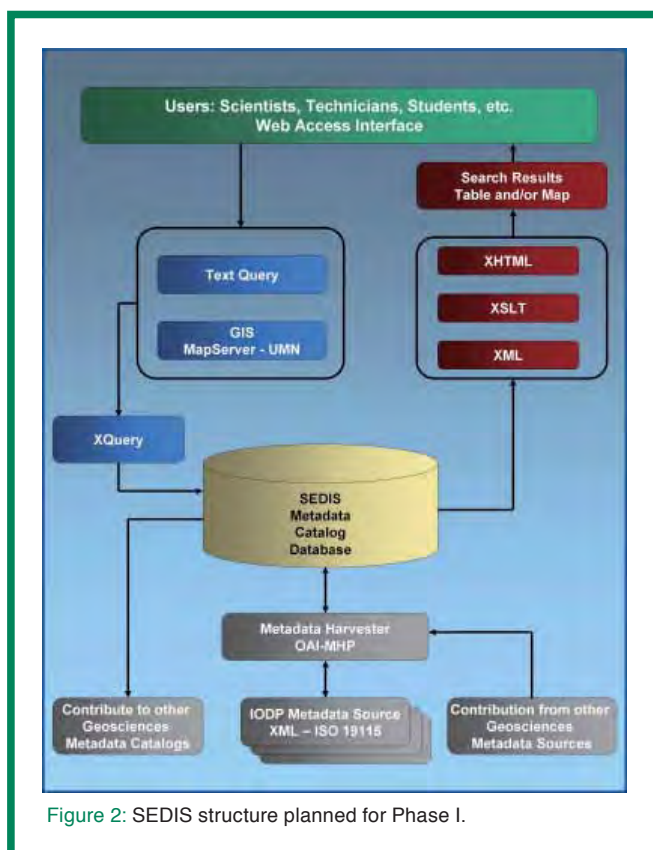Figure 1: IODP data and metadata flow for SEDIS.

Figure 2: SEDIS structure planned for Phase I.

be aimed at the broadest possible interoperability of both data and related data manipulation and visualization tools. International standards and open-source components will be used whenever possible. The development of SEDIS will take place in three phases during 2006–2008. Phase I will concentrate on creating a searchable catalog of all IODP and ODP drilling and logging data stored in the primary databases maintained by the IODP operators. A competitive request for proposals is planned for the second quarter of 2006, and implementation will happen in early 2007.

SEDIS Phase I will allow the user to discover, via database search queries and geographic information system (GIS) techniques, what type of measurements are available for specific drill sites and holes. It will also provide links to the data itself, independently of acquisition, storage location or archival status (Fig. 1). This concept allows the data to be stored in distributed systems while offering centralized access. The use of internationally recognized metadata standards will also allow other geoscience portals to harvest IODP metadata and make it more easily visible and available to researchers outside the IODP community.

SEDIS Phase II is intended to closely tie together the program data, program publications and post-drilling data from core samples, the boreholes, and any related experiments. SEDIS will not store publications but will have the capability to do searches based on keywords and Digital Object Identifiers (DOI) for the required information within the underlying program databases. Most IODP post-drilling research will be published in the open peer-reviewed litera-

ture, and the system will also need to access as broadly as possible all program data associated with such publications.

In SEDIS Phase III, advanced tools enabling users to search, extract, and combine data as well as to integrate advanced data visualization and mapping tools are to be developed. Implementation likely will involve other organizations and projects that have developed certain specialized applications (e.g. CHRONOS, CoreWall, Stratigraphy.net and others). The goal is to enable the user to extract any part or combination of data from any drilling expedition, and to provide methods to process these data with online viewing, mapping, and manipulation tools from a single entry point. Figure 2 provides an overview of the SEDIS structure.

Core Samples, Drilling Proposals and Data in Support of Drilling,

There are three IODP core repositories: Bremen (Germany), College Station (Texas, U.S.A.) and Kochi (Japan). To facilitate core sampling and tracking sample history and sample availability, an IODP Sample Materials Curation System (SMCS) is targeted for implementation during 2007. The SMCS will provide single Web-based access for sample requests and sample management. Sample requests will be submitted through the SMCS frontend to the appropriate curator and recorded in the SMCS database for user and curator reference. It is to be developed in close collaboration with the drilling and IODP core repository operators. Information about physical samples in the SMCS will eventually be linked to sections within the expedition publications and the metadata in the SEDIS via DOI, interlinking all the information.

In 2005 a new proposal database for IODP drilling proposals and a new fully electronic Site Survey DataBank (SSDB) were initiated. Both projects move proposal and data submission to an all-digital stage, making it easier to upload, access, and manage proposal information and site-survey data. The proposal database is still under development and is expected to be online by mid-2006. The SSDB is in operation, but new functionality is still being added (see article on page 40 in this journal).

Authors

Bernard Miville, Emanuel Soeding, and **Hans Christian Larsen**, IODP-MI, CRIS Building, Rm. 05-101, Hokkaido University, N21 W10 Kita-ku, Sapporo, 001-0021 Japan, e-mail: bmiville@iodp-mi-sapporo.org

Related Web Links

http://sedis.iodp.org
http://www.iodp.org/data-management/