

Big data management in ocean science

*Jota Kanda¹

1. Department of Ocean Sciences, Tokyo University of Marine Science and Technology

In the field of ocean science, an international framework for sharing observation data has been established from an early stage. The Intergovernmental Oceanographic Commission (IOC) was established in 1960 at UNESCO primarily for international cooperation in the field of ocean science. A mechanism for exchanging observation data has been established immediately after its foundation. Specifically, each member country has a national data center, a system for collecting, managing and providing ocean data in each country. International data exchange is conducted between these national centers. In Japan, the Hydrographic and Oceanographic Department of the Japan Coast Guard operates the Japan Oceanographic Data Center (JODC). The IOC also has the International Oceanographic Data and Information Exchange (IODE) program, which works on technical considerations and future directions for data exchange and disclosure. In this way, the open oceanographic data has a relatively long history, and a system has been built up with accumulating experience. Meanwhile, ocean science and observation technology have advanced remarkably, and the scope of activities such as the IOC has diversified. New aspects have also become prominent in the exchange of oceanographic data. Wide-area and continuous observation data from satellites and floats have become available in addition to the observations made by traditional vessels. While the amount of data becomes enormous, observing data in real time with improved data processing capabilities has enabled data assimilation into model calculations and eventually accurate predictions of ocean status. One of the major directions is to deal with the huge size data in real-time. The Argo Project, as an example, is a global scientific program where international organizations such as the World Meteorological Organization (WMO) and the IOC and research and administrative organizations in various countries cooperate to deploy the Argo floats that automatically ascend and descend to observe. The vast amount of data obtained from many floats of various countries is released free of charge, released in real time and archived through quality checks in parallel. Observation data used to be passively collected from individual researchers/research-groups and government agencies with their own purposes. In contrast, large-scale international collaborative projects with a focus on the relationship between climate change and the ocean have been in progress and produce enormous data sets. In consideration of the needs for various oceanographic data, frameworks such as the Global Oceanographic Observation System (GOOS) are established in order to coordinate the overall oceanographic observations by rational input of observational resources and by examination of observational designs. With the cooperation of several members of the Oceanographic Society of Japan, I will discuss the current status, issues, and future direction of data management in ocean science.

Keywords: Ocean, Open data, Big data