

China Contribution to JTC 1 Study Group on DCOMP

In the information age, digital information resource has become a national strategic resource. Mass of digital information resources are being generated while many of them are missing. With a positive response, countries all around the world carry out many projects for long-term preservation of digital information in order to record and inheritance their culture and civilization in digital form. However, long-term preservation technology for digital information developing rapidly and the disunited standards, seriously affect the sharing and exchange of digital information. OAIS reference model provides a conceptual framework aiming at the long-term preservation and maintenance of digital content. It describes the environment and function organization of the archives system which is supported by infrastructure of the archives information management. As a benchmark framework for long-term preservation of digital information, it has been widely accepted by many national instructions and the saving sector from all over the world. Then it has been applied to many practical projects in order to preserve digital information.

OAIS reference model describes the entire process of archives information saving function, including processing, archives storing, data managing, accessing, releasing. OAIS system meets the basic needs of long-term preservation of digital content from the perspective of the system function. But explicit data model is required when long-term preservation of digital content is actually implemented. Meantime, data model should meet the whole lifecycle management need of long-term preservation of digital content.

Data is the core object of digital content management. It is necessary to establish a unified data model in order to support the long-term preservation of data content. There are some requirements from the perspective of data as follows:

- Security and access control of digital content
- Low cost migration of digital content
- Understanding digital content need resource tools to support
- Accessing digital content require to back to the historical data
- Ensuring the quality of service when the digital content distribution

The problems arising from these needs are very complex, we utilize hierarchical structure to establish data model. This data layered architecture converts the entire large and complex problem into several sub-problems which are easy to be solved, with good expansibility. It can meet the needs of long-term preservation of digital content and the whole lifecycle management.

With the widespread use of information technology, multimedia data grows more rapidly than others. Many multimedia data do not consider the need of their long term preservation when they were created. The users must consider the issue that long-term preservation, management and usage of multimedia data.

Multimedia data which has rich content is an important part of data resources. There are large amount of encoding and file formats, some of them use the international standard coding format, and some utilize proprietary compression format which is defined by company or community, even the international standard format, there are also many species. Lack of good compatibility is the status of multimedia digital content management and protection.

The data content of each format has its unique value and history, while we hope that these resources would be durable, we hope that information technology would achieve this aspiration.

But in fact, not every data formats are all suitable for long-term preservation, sometimes new technology develop rapidly, while the old technology was accelerated to die out so that some form of data content eliminate faster. Therefore, we should choose some data format which can meet the need of long-term preservation, maintain them for long-term under the guidance of the standard, and research on the conversion of data format, to preserve variety of data content as far as possible.

In summary, the China national body recommends that the JTC 1 Study Group on DCOMP focus on following area:

1. Digital Content Management Data Model

- a. The establishment of framework of layered data model and the definitions of each layer of data model.

- b. The relationship among the layers and interface specification in the model, and the interaction between layers and the operational primitive definition.

- c. The framework implementation of data model.

2. Multimedia Data Content

- a. The requirements of multimedia digital content format for long-term preservation, establishing criteria of file format evaluation, maintaining a standard format that is adapt to long-term preservation.

- b. The technical framework, conversion protocol, interface specification and resource tool for multimedia data conversion in order to achieve data transfer and sharing.