Mapping of REKIHAKU Database’s Fields to Common Metadata for Integrated Retrieval

ADACHI Fumio

The National Institutes for the Humanities established an integrated retrieval system for the over 100 diverse databases provided by its member organizations and launched the service in April 2008. This system is characterized by its intuitive usability and enhanced algorithm to allow users to retrieve any data regardless of its operation method or location. Like most other search engines, the system enables users to retrieve information matching their search terms and shows a list and detailed description of search results. In this system, however, a common metadata set plays an important role in order to retrieve data from databases developed separately and therefore organized by different attributes. A key issue of realizing the system is how to map data attributes from the different databases to this common metadata set. This paper describes a basic idea of mapping 43 databases of the National Museum of Japanese History for this integrated retrieval system.

In order to show detailed search results in an easy-to-understand way, it is essential to map data attributes in a way that some metadata searches do not make too many requests. Based on this idea, the Dublin Core (DC) Metadata Elements, adopted as a common metadata set, is not highly compatible to search engines retrieving information from databases based on descriptions of physical materials because the meanings of the elements need to be expanded. This problem has been revealed by comparing the Museum Core Metadata Elements (which were developed to share information on the collections of the Museum) with the DC Metadata Elements. The Museum Core Metadata Elements can support the mapping of databases not only of physical materials but also on historical sites, as well as various research documents and their research purposes. This is because the Museum Core Metadata Elements can be used as description elements to retrieve information on people, time, and space, regardless of content in the form of role and achievements, and because they are often the interchangeable elements of name, type, and subject that attribute specific features to the objects being described. If there is a need to use the metadata set for more diverse purposes, it will be essential to abstract the remaining elements and explore their similarities.

Key words: transversal retrieval, resource sharing, Dublin Core meta-data, museum materials, information of humanities