

Research on Metadata Standards of Biomedical Data Repository

In recent years, scientific data has become another important scientific research resource after scientific research literature. Scientific metadata plays an important role in the field of management, long-term preservation, sharing and open access of scientific data. However, a large number of diversified metadata standards and widely different metadata elements have become an obstacle to the sharing of medical data resources. To solve this problem, the first step of this study is to establish a comparative analysis framework for some authoritative metadata standards in the biomedical field, such as Darwin Core, Data Tag Suite (DATS), metadata standard of Dryad, and some frequently-used general metadata standards, such as DataCite Metadata Schema, Dublin Core, metadata standard of re3data. The content of the analysis framework includes the construction goals, characteristics, core elements, the categories of the described objects, the scope of application, etc. of each metadata standard. Then by comparing the common elements and similar elements of metadata standards, we discuss how to realize the interoperability among different metadata standards, and how to achieve scientific data integration and unified retrieval among different systems through metadata mapping, reusing and sharing. Through this research method, the scientific data of different regions, languages and structures can be integrated together organically and meet the needs of people more effectively and systematically. At the same time, providing a theoretical reference for the sharing and management of biomedical data resources in China.

Keywords: scientific data; metadata standard; biomedicine