A Study on the Content Quality Management Procedures in the OAK Portal System

Jong-Myung Baek*, Hye-Kyong Hwang**

*,**KISTI, South Korea

E-mail : jm100@kisti.re.kr*, hkhwang@kisti.re.kr**

1. Introduction

A repository becomes most ideal for users of an institute in case their research outcomes may be deposited in the repository through self-archiving and in case a registered data service may be used free of charge. Meanwhile, for various reasons, low quality data has been accumulated and serviced in the repositories of Korean institutes. This study introduced the Open Access Korea (OAK) Portal, which is a representative Open Access (OA) contents collection and offering service in Korea, and proposed the elements of content quality management to be considered during collection and a workflow. Also, by developing and applying this to the system, the study aimed for the enhancement of OA system contents quality that will be offered in the future.

2. OAK(Open Access Korea) Portal

OAK(Open Access Korea) Portal is an information service developed by KISTI in Korea.

It is possible to carry out an integrated search by harvesting open access contents metadata of an institute that has installed an OAK repository that is a D-Space-based Korean type repository. The OAK Portal service offers integrated search by harvesting open access contents that have been produced and deposited by researchers who belong to a institution using the OAI-PMH protocol, and also it has furnished relevant information and trends for Korean Open Access.

3. Quality issue of contents harvested in OAK Portal

In accordance with the basic idea of Open Access, it may be most ideal that users belonging to an institute register their own research outcomes through self-archiving and that such registered data may be offered so that it is used free of charge; however, it has been shown by research from Baek Jong-myung and Lee Eun-jee(2013) that for diverse reasons there are various types of quality issues with the contents harvested on the OAK Portal. The first problem is that only meta data exists and no original contents exist. The second problem is that it is difficult to distinguish the type of certain data in the OAK Portal as there are so many diverse "types" within the meta data.

Such a problem was caused by various reasons; however, it was discovered that the main causes are an overlapping archiving problem with the legacy system and copyright issues Also, for Korean repositories, in most cases, large data is registered through a repository manager (or a library manager) rather than through self-archiving, and voluntary registration of good data has been very rare.

4. Contents quality management system of the OAK Portal

As with the OAK Portal, a system that collects and services contents in various repositories needs a workflow on which only quality contents are serviced selectively by conducting contents quality management independently. In particular, such independent quality management is even more necessary in case of collecting diverse types of unstructured data.

This study proposed a quality management workflow (Fig. 1) based on these results, and implemented such a system. In this system, a manager defines quality management meta items as essential/selective items, and it is possible to conduct a regular test of items within the OAK Portal in reference to defined meta items. Through this, a system is developed in which a manager recollects relevant items in an OAK repository by institute by identifying and determining error data.

5. Detailed function

The key functions for quality management are made up largely of four items. First is the management of the meta field automatically harvested by a collection scheduler. It is a function with which a meta field can be added by a manager, and contents can be selected and serviced by designating essential items. Second is a "harvested contents confirmation function" with which harvested contents can be approved or rejected by a manager by confirming the harvested contents and looking at the original copy. Third is a function of showing the quality management history of a manager as a history function. Finally, fourth is a "service data type function" suitable for 8 types that is able to select and save types that are designated differently by each institute; thus, it can compose the data as consistent data types.
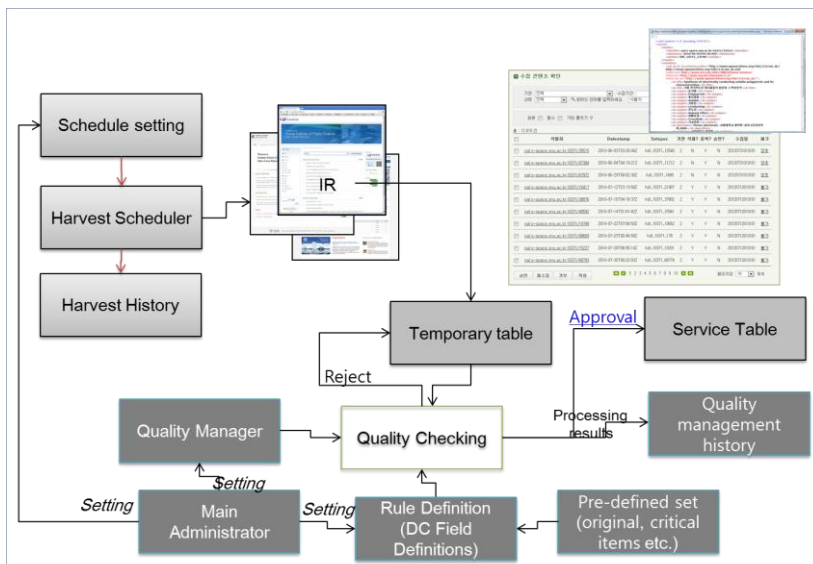

Figure 1. Workflow of Quality management function

6. Conclusion

The process developed and applied in this study may help to improve contents quality in the OAK Portal; however, quality improvement within a repository is fundamentally even more important. Also, it is crucial for a manager to define meta data using the established Dublin Core meta data terms, combine this with reclassified data type, identify the original text type (automatically or manually), and hand over only quality contents to a service table.

With the accumulation of standardized quality contents through consistent and diverse investigations and analyses, it is hoped that Open Access contents will be extended by participation in the Korean academic communication market.

7. References

[1] J. M. Baek, E. j. Lee, "A Study on the Contents of Institutional Repositories in Korea", Korea Society for Information Management Conference, 2013.
[2] Hwang, Hye-Kyong, Lee, Jee-Yeon, "Analyzing the factors affecting the successful deployment of the open source based institutional repositories". Journal of Korea Society for Information Management, 26(4), pp 35-57