Rekisuke: a system for sharing bibliographic information of historical documents through Interdisciplinary collaboration

*Mika Ichino[1,2], Yukie Hashimoto, Kooiti Masuda[3], Asanobu Kitamoto[2,1], Junpei Hirano[4]

1. Center for Open Data in the Humanities, Joint Support-Center for Data Science Research, Research Organization of Information and Systems , 2. National Institute of Informatics, 3. Tokyo Metropolitan University, 4. Faculty of Liberal Arts, Teikyo University

The records of numerous historical documents in Japan have been used for research in diverse scientific fields ranging from seismology, climatology and astronomy, to humanities, social sciences and historical studies. Unfortunately, when processing and applying information found in historical documents for the purpose of scientific research, the one may encounter certain difficulties despite the instrumental observations and the born digital data which can be accessed and acquired easily. Promoting in interdisciplinary fields is one of the measures used to solve these issues related to historical documents. Even under these circumstances, the cooperation with researchers from various fields relaying on historical documents has started. It will develop into interdisciplinary collaborative research to facilitate understanding of spacio-temporal association between nature and human society with a potential to become a novel scientific field.

The processes not only include the finding, obtaining, transcribing and reading the descriptions in the documents, but also structuring, analyzing and integrating, and sharing the data. Transcribing is most difficult for scientists and engineers, because they are not always familiar with historical documents. In the first place, finding the historical documents relevant for research is not always an easy task. It might be a real challenge to identify the best one that suits our needs.

Until now, the research studies using historical documents have been progressing in separate fields, but sometimes researchers extract data from the same materials. For example, Diaries contains not only daily weather conditions, but also important records such as meteorite and earthquake events. Over the years, individual research groups have extracted and digitized the data necessary for their research field only. Sharing that data is more difficult than the information about their existence, due to the secrecy of research and intellectual property legislation.

Hence, we are focused on developing "Rekisuke", which is a system for interdisciplinary sharing of the variety of information, experience and knowledge obtained from each individual research field. Moreover, any improvements of increasing usability, linking sources, displaying the results, user interface, and increasing the data number towards the release are in great need.

The system can thus help researchers to avoid repeating the same processes/tasks that have been already done by others. An important role of the system is to provide information about the existence and the status of the various material and data. Users can help each other in during the different phases of the process, not only finding and obtaining the data but also reading and structuring them, because it can contain descriptions of various information, such as the state of the materials (paper material, image, transcribed text data, published book, digital text data, structuring data, etc.).

It has some functions enabling the search based on particular keywords, such as weather, earthquake. One of them is a notable function to search by temporal and spatial range by using APIs developed for historical temporal and spatial. It is also helpful for libraries which need to promote their historical documents. We started the collaboration with a local library for promotion of their documents and utilization of our system.

Keywords: Historical documents, Sharing information and data, Interdisciplinary collaboration, Metadata, Paleo climatology