Suppl 2. Interviewee's comments.

1. Transition from Individual Research in the Humanities to Digital Scholarship through Interdisciplinary Collaboration

Conflicts between data and technical specialists sometimes occur due to a lack of clearly defined responsibilities.

In this system, specialists in different fields conduct research with each other. I believe it is better for the collaboration if each member focuses on his or her area of specialization (P3).

2. Converting Analog Humanities Resources into Data

Current research teams need to transcribe the images manually, which is both time-consuming and costly.

Humanities Korea Research Project (HK) research projects have masters and doctoral students type out analog texts because it is difficult to convert the images to digital data using OCR tools. Another option is outsourcing the conversion, which is expensive. Transcription fees constitute a significant portion of DH research costs, and this is a problem not only for researchers currently working on the HK project, but also for those who will undertake future research (P4).

3. Support for Data Modeling and Analysis Training

I would like technical training, in which the rules of data analysis are used in spreadsheets, to be held as specialized lectures for each generation of researchers (P1).

When workshops on related word searches, frequency searches, co-occurring word analysis using big data in terms of newspaper texts were provided, along with opportunities for practicing during semester breaks, many researchers who needed re-training on humanities were highly interested and participated (P3).

The literacy training support service of both academic and national libraries provides digital humanities resources as well as a training support services that focus on imparting practical technology training.

4. Humanities Data Curation

As an anthropologist, I want to find ways to accumulate daily life experiences in Korea and make my hard-found data meaningful and worthwhile (P6).

Databases of literary works constructed by processing data were not public open and could not be used.

The literary works from the Institute for the Translation of Korean Classics were constructed as data, but significant expenses were spent on tagging the names of people and places. Although we should be able to download all of the tagged data, to answer the question of "Can we see the flow of thoughts over the ages?", a limited amount of data was distributed (P2).

Some have also argued that a Korean AI translation system should be built to empower future generations of humanities researchers to access classical humanities resources discovered through curation and translated into modern languages.

5. International Collaboration to Expand Digital Scholarship Research

Openness, sharing, and collaboration are important factors of DS, and the fact that related infrastructure is not shared well among institutions seems to be a problem (P4).

Humanities data portal services that provide information on the location of digital humanities resources in Korea, as well as offer detailed information about copyrights, should be planned for and operated. Additionally, networks need to be formed between overseas institutions and researchers, and it is necessary to use internationally standardized frameworks for data resources to be shared as a foundation.

Although Korea's digitization background and databases have their advantages, they need to adopt international sharable data architecture and sharing methodology; an interoperability framework platform is needed (P5).

In Korea, although DS provides original text in the form of XML with an appropriate ontology, it needs a circulation platform based on an international sharing standard. For instance, an interoperability framework—such as IIF (https://iiif.io/), which is a data construction methodology for images—should be considered for annotating images and to share, open up, and disseminate resources in a multinational environment.