Research information service development plan based on an analysis of the digital scholarship lifecycle experience of humanities scholars in Korea: a qualitative study

Jungyeoun Lee¹, Eunkyung Chung²

¹Department of Library and Information Science, Joongbu University, Geumsan; ²Department of Library and Information Science, Ewha Womans University, Seoul, Korea

Abstract

Purpose: Given the impact of information technologies, the research environment for humanities scholars is transforming into digital scholarship. This study presents a foundational investigation for developing digital scholarship (DS) research support services. It also proposes a plan for sustainable information services through examining the current status of DS in Korea, as well as accessing, processing, implementing, disseminating, and preserving interdisciplinary digital data.

Methods: Qualitative interview data were collected from September 7 to 11, 2020. The interviews were conducted with scholars at the research director level who had participated in the DS research project in Korea. Data were coded using Nvivo 14, and cross-analysis was performed among researchers to extract central nodes and derive service elements.

Results: This study divided DS into five stages: research plan, research implementation, publishing results, dissemination of research results, and preservation and reuse. This paper also presents the library DS information services required for each stage. The characteristic features of the DS research cycle are the importance of collaboration, converting analog resources to data, data modeling and technical support for the analysis process, humanities data curation, drafting a research data management plan, and international collaboration.

Conclusion: Libraries should develop services based on open science and data management plan policies. Examples include a DS project liaison service, data management, datafication, digital publication repositories, a digital preservation plan, and a web archiving service. Data sharing for humanities research resources made possible through international collaboration will contribute to the expansion of new digital culture research.

Keywords
Digital humanities; Digital scholarship; Digital scholarship research life cycle; Digital scholarship service

This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Introduction

Background and rationale
Digital scholarship (DS) seeks to obtain innovative and collaborative results through the digital curation of resources, the multidisciplinary nature of research, and sharing research processes and outcomes. DS refers to the use of digital evidence and methods, digital authoring, digital publishing, digital curation and preservation, and digital usage and reuse of scholarship [1]. The definition of DS is clarified by its associated activities, which include text and data mining, machine learning, data visualization, digital mapping, and digital editions [2,3]. The Association of Research Libraries also offers research support services for DS [4,5]. Some examples of DS research include Slave Voyages (https://www.slavevoyages.org), Mapping the Republic of Letters (http://republicofletters.stanford.edu), and the Perseus Project (http://www.perseus.tufts.edu/hopper/).

In Korea, institutional database construction projects began as part of a national digitization project for analog humanities resources, starting in the late 1990s and running through 2010. A vast number of institutional records was digitized. However, DS in Korea can be succinctly in digitized databases. The conversion of analog resources into data and subsequent analysis have not yet been performed adequately or fully and, furthermore, DS research support services are rarely provided [6–8].

Objectives
By analyzing DS lifecycle research, this study aims to propose a DS service strategy. This is achieved by conducting foundational research on the characteristics of DS research services requisite for each research stage. The analysis focuses on openness, interoperability, usefulness, reusability, integration, and participation in the DS information service within libraries.

Methods

Ethics statement
The interview data collected in this study were recorded after obtaining consent, in compliance with research ethics regarding personal information protection and the use of data for research purposes.

Study design
This was a qualitative study based on interviews. It was described according to Standards for Reporting Qualitative Research [9].

Qualitative approach and research paradigm
A representation of the entire DS research cycle was developed based on Brügger’s schematic representation of the research process (Fig. 1) [10]. In response to the changes in analog, digitized, and born-digital materials, DS experts divided the re-
Research information service development plan

search programs according to access, analysis, discussion, and the dissemination of data into research planning, execution, publication of results, dissemination, preservation, and the reuse of research findings.

**Researcher characteristics and reflexivity**
The researchers are experts in library and information science with over 18 years of experience.

**Context**
Interviews were conducted in a question-and-answer format based on a semi-structured questionnaire (Suppl. 1), and the data were analyzed using semantic unit coding and clustering.

**Sampling strategy**
The participants in this study were researchers with experience equivalent to or greater than principal investigators who had planned and established national database projects in the humanities and obtained research results in DS.

**Data collection methods**
Data were collected through interviews. The interviewer experts also collected empirical data about data collection, analysis, curation, and reuse within the DS research program and then conducted data analysis.

**Data collection instrument and technologies**
Face-to-face online interviews were conducted, with each interview lasting an average of 1 hour and 7 minutes (Table 1). The voice recordings were transcribed into written documents and used as the primary data for the study.

**Units of the study**
The interviews took place from September 7 to September 11, 2020. The participants’ academic fields included history, Korean literature, and cultural anthropology. The participants held doctoral degrees and were professors and researchers at universities and government agencies.

**Data processing and analysis**
The interviews were transcribed, and responses were categorized by theme using Nvivo 14 (Lumivero). Content analysis was performed by creating group clusters while coding to restructure relevant theme nodes (Suppl. 2). To evaluate the study’s reliability and validity, cross-analysis among researchers was conducted. Based on the results of coding performed by two coders, intercoder reliability was measured using Cohen’s κ statistic and was found to be 0.838, which falls within the range of substantial reliability [11].

**Techniques to enhance trustworthiness**
No further process was implemented.

**Results**

**Synthesis**
Transition from individual research in the humanities to DS through interdisciplinary collaboration

The central nodes of the DS research cycle comprise five stages: research planning, implementation, publication of results, dissemination, and the preservation and reuse of research data (Fig. 2, black rectangles). Additionally, there were 32 subnodes within the central nodes. “Collaboration” was the central topic throughout all the stages. As shown in Fig. 2, the “collaboration” node was extracted from the research planning and data preservation and reuse processes because the researcher considered it to be the most significant characteristic of DS (Fig. 2, red circles).

First, collaboration is needed for digitizing research data in the research plan stage. Second, cooperating with data analysts and information managers is required for converting the research implementation stage into data and technical analysis. Third, technical collaboration is essential for publishing results online, in addition to publishing in academic journals. Fourth, DS institution networking is vital to facilitate research dissemination. Fifth, working with data management institutions is needed for sharing data for preservation and reuse.

**Table 1.** Interviewee backgrounds, interview modality, and interview duration

<table>
<thead>
<tr>
<th>ID</th>
<th>Research discipline</th>
<th>Affiliation type</th>
<th>Gender</th>
<th>Position</th>
<th>Degree</th>
<th>Interview duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>History</td>
<td>University</td>
<td>Woman</td>
<td>Professor</td>
<td>PhD</td>
<td>50 min</td>
</tr>
<tr>
<td>P2</td>
<td>History</td>
<td>Government institute</td>
<td>Man</td>
<td>Researcher</td>
<td>PhD</td>
<td>1 hr and 11 min</td>
</tr>
<tr>
<td>P3</td>
<td>Korean literature</td>
<td>University</td>
<td>Man</td>
<td>Professor</td>
<td>PhD</td>
<td>1 hr and 20 min</td>
</tr>
<tr>
<td>P4</td>
<td>Korean literature</td>
<td>University</td>
<td>Man</td>
<td>Professor</td>
<td>PhD</td>
<td>45 min</td>
</tr>
<tr>
<td>P5</td>
<td>Korean studies</td>
<td>University</td>
<td>Man</td>
<td>Professor</td>
<td>PhD</td>
<td>1 hr and 15 min</td>
</tr>
<tr>
<td>P6</td>
<td>Cultural anthropology</td>
<td>University</td>
<td>Man</td>
<td>Professor</td>
<td>PhD</td>
<td>1 hr and 5 min</td>
</tr>
</tbody>
</table>
However, difficulties arise in finding and collaborating with researchers from each academic field during each stage of the DS project processes. A liaison service for DS project management is necessary, such as public service librarians supporting the organization of an interdisciplinary research team or technical services librarians consulting during the conceptualization, planning, and implementation stages where interdisciplinary collaboration is required [12].

Converting analog humanities resources into data

Humanities scholars access and interpret information through answering academic questions based on data. The first DS stage is converting original analog forms of cultural heritage, such as ancient documents, maps, or records, into digital resources. The basic activities of digital humanities (DH) scholarship involve either making analog materials, including maps, manuscripts, archaeological remains, specific editions, or primary documents of historical or recent events, available in a digital format or creating and using born-digital materials. Converting information from one form into another is the activity of remediation. This preserves information in a file format, which a computer can then process. DH projects often engage with both original analog materials and born-digital materials, but all are managed in the form of digital assets [13].

In Korea, national digital database projects have been conducted at institutions that have old historical resources. However, the problem is that most of the databases were built by scanning original analog texts in the form of images. These digital images should be converted to data for DS research. This is the case because nationally supported database con-
structure projects were conducted as job-creating initiatives for the digitization of analog resources. Moreover, the focus was on digital preservation during the IMF (International Monetary Fund) economic crisis in Korea. As a result, a standardized data structure in which humanities resources can be read and interpreted as data was not provided from an information usage perspective. Therefore, to support DS research, technical support services are needed for designing a standard data structure and converting analog and image texts into data.

Support for data modeling and analysis training
For humanities data to be analyzed according to research purposes, data modeling and analytic technologies should be used. Data modeling refers to datafaction and modeling. This process involves abstracting discrete values from a phenomenon or artifact. These values may be expressed either in numbers or texts and are necessarily a reduction of complex materials into a computational form. With data, one can automate processes, such as sorting, counting, comparing, or making statistical assessments. Materials or phenomena of almost any kind can be turned into quantifiable or discrete data [13,14].

Of the six participants in this study, four had personally performed data modeling and analysis tasks. However, each researcher employed a different data processing method in their collaborative research, resulting in difficulties like unbalanced datasets. Although research data should be modeled, constructed, and analyzed according to their characteristics, there is hardly any one program in which the process can be studied systematically. Therefore, libraries need training support on the data characteristics most suitable for DS research, modeling, data construction guidelines, and data analysis technologies. For instance, curricula of various levels on the analysis of social networks based on data, texts, and using databases should be developed. Furthermore, it is necessary to provide support to facilitate collaborative research among participants by forming networks.

Humanities data curation
In DS research, the crucial point is not to obtain hundreds of gigabytes of rapidly increasing data, but rather to select valuable resources. Researchers need the development, creation, and structuring processes of qualitative resources worthy of cultural studies and emphasize the need for qualitative data management of resources, such as the origination of metadata, whose context can be interpreted for the reuse of data.

Data curation has been defined as the active and ongoing management of data throughout its entire lifecycle of interest and usefulness to scholarship. Curation has also been closely associated with DH practice, in addition to digitally preserving and curating cultural heritage material. It includes the following activities and processes: description (documenting the context and relationship of various forms of research data), annotation (enhanced information on the data with more granularity and context), collection and aggregation (connecting data and teams), storage (maintaining a platform for stable and accessible data), and migration (to ensure continued access via emulation or preservation) [14,15]. The DS research projects undertaken by this study’s participants were not conducted in collaboration with libraries from the data curation process. Instead, the research team collected and processed data themselves.

In the future, a foundation services should be established so that data curation, which can support research projects through collaboration among libraries at the beginning of the research project, can be conducted collaboratively early on. It is essential to describe and annotate the context of data and to develop and provide services where such data can be used collaboratively between librarians and researchers.

Sustainable research data management policy
Open science, which is based on sharing and collaborating on the latest research data, drives the policies and trends of global research fields. Depending on the characteristics of research areas, data sharing policies and data management plans (DMPs) should be submitted before research projects begin. Recently, the National Endowment for the Humanities, which supports DS research, has started requesting planning, preservation, and management plans for research data at the preparation stage for DS projects. This follows the National Science Foundation’s “Managing and Sustaining the Project Assets” guideline [16].

The importance of research data, which originated in scientific and technological fields, is now expanding to support humanities and social science work. The goal is for data to be managed, preserved, and reused by third parties during the curation process; that is, the aim is to extend beyond just managing research data. While data management refers to the direct creation and use of data by researchers during the data lifecycle’s active stages, curation ensures that data can continue to be managed both during and after research.

In Korea, services for DS projects often face the threat of discontinuation because maintaining web publications on servers becomes difficult given the lack of funding after the research ends. Unfortunately, this leads to the loss of national research and cultural resources, as well as the inability to reuse data in follow-up studies. To address this issue, the infrastructure for preserving and reusing data must be strengthened by expanding digital preservation strategies and digital archives as large as data centers in libraries for DS work.
International collaboration to expand DS research

DS is based on openness, sharing, and collaboration in research. Furthermore, the dissemination, sharing, and openness of research results in the humanities, in addition to scientific and technological research, are gradually expanding. The majority of DS research in Korea, which has abundant cultural resources, originates from data derived from the historical relationships among Korea, China, and Japan. Conducting research on and providing open access to data, and sharing results are the starting points for disseminating human cultural heritage research in the future.

Discussion

Key results

DS should be able to datafy, which is not just a matter of digitizing, but also transforming into an analyzable format, related resources and interpret them using humanities research questions. This study found that although DS research has remained at the database construction stage, it is expanding into individual or team research. However, researchers have experienced difficulties in collecting, managing, applying, and analyzing data because of the lack of library-based, collaborative research support services. This has been the case regarding research data curation, data management, and analysis processes, all the way from the early to the final stages of DS research. Therefore, this study analyzed the DS research cycle, explained the five stages of the research process, and proposed library services needed at each stage (Fig. 3).

First, the most significant characteristic of DS research is “collaboration.” This is because DS interprets humanities questions from various viewpoints based on humanities data. In addition to humanities researchers and those working in other disciplines, data management, and analysis specialists collaborate to derive results. Second, because DS research is based on humanities data, data curation and management policies, as well as data management such as data collection, management, analysis, and preservation, are essential, and the role of the library is crucial. Third, the following stages of DS research were identified: research plan, research implementation, publication results, dissemination of research results, and the preservation and reuse of research data. Fourth, the library information services required at each stage are as follows. For the research planning stage, there is a need for DH conceptualization support services and DS project liaison services. In the implementation phase, technical support and training services, such as data curation and management, datafication, and data modeling, are necessary.

In the publication stage, content management system support should be provided to researchers for submitting and publish-
ing their work, and for publishing their results on the web. Additionally, support should be given for operating a content management system. In the dissemination phase, support for collaboration should be offered by constructing an international standard framework support service and a network for sharing domestic and international DH resources to share data internationally. In the preservation and reuse stage, data management, long-term hosting, and web archiving services should be included in digital preservation plans.

As such, libraries should not limit themselves to constructing and providing digital resources. Rather, they need to continue opening up resources and sharing services, such as data structuralization, technology analysis, training program development, data management, web archiving services, and international collaboration.

Interpretation
If libraries actively engage in collaboration processes with reference to the service strategies for each stage of DS research proposed in this study, establish curation strategies from project planning, and implement management through the DMP process, then DS research data will be preserved for an extended period and reused by future generations in their studies. As such, digital cultural research will continue.

Limitations
This study has conducted an empirical investigation with a limited number of researchers who have experience in DS research. If expanded to include librarians and other stakeholders for the application of DS library services, a more detailed service improvement plan can be proposed.

Conclusion
DS is a convergent field of research based on open science. If libraries serve as the center of DS collaborative research, they can truly empower future generations to reuse research data and maintain the sustainability of DS research resources. This could be done via research support throughout the DS research cycle, as well as research data management, data curation, and web archiving.

Conflict of Interest
No potential conflict of interest relevant to this article was reported.

Funding
This work was supported by a Joongbu University Research and Development Fund in 2022.

Data Availability
Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Supplementary Materials
Supplementary materials are available from https://doi.org/10.6087/kcse.309.

Suppl. 1. Interview questionnaire on research experience based on digital scholarship research lifecycle.
Suppl. 2. Interviewee's comments.

References


