Chih-Ming Chen* and Ming-Yueh Tsay "Applications of Collaborative Annotation System in Digital Curation, Crowdsourcing, and Digital Humanities", *The Electronic Library*, vol. 35, issue 6, 1122-1140, 2017. **[SSCI 收錄] (post print version)**

Applications of Collaborative Annotation System in Digital Curation, Crowdsourcing, and Digital Humanities

Chih-Ming Chen[1*], Ming-Yueh Tsay [2]

[1] Graduate Institute of Library, Information and Archival Studies, National Chengchi University, Taipei City, Taiwan, ROC; E-mail: chencm@nccu.edu.tw

[2] Graduate Institute of Library, Information and Archival Studies, National Chengchi University, Taipei City, Taiwan, ROC; E-mail: mytsay@nccu.edu.tw

[*] Corresponding author

Type of paper: Technical paper

Keywords: Collaborative annotation system, Digital curation, Crowdsourcing, Digital humanities

Abstract—Collaboratively annotating digital texts allows users adding valued information, sharing ideas, and creating knowledge. Most importantly, annotated content can help users obtain a deeper and broader understanding in a text compared to digital content without annotations. This work thus proposes a novel collaborative annotation system (CAS) with four types of multimedia annotations including text annotation, picture annotation, voice annotation, and video annotation which can embed with any HTML web pages to enable users to collaboratively add and manage annotations on HTML web pages and provide a shared mechanism for discussing about shared annotations among multiple users. By applying the CAS in mashup on static HTML web pages, this study aims to discuss the applications of CAS in digital curation, crowdsourcing, and digital humanities because of existing strong relations among them.

Design/methodology/approach—This work adopted Asynchronous JavaScript (Ajax) and Model-View-Controller (MVC) framework to implement a CAS with the reading annotation tools for knowledge creating, archiving, and sharing services as well as applied the implemented CAS to support digital curation, crowdsourcing, and digital humanities. A questionnaire survey method was used to investigate the ideas and satisfaction of visitors who attended the digital curation with CAS support in the item dimensions of the interactivity with displayed products, the attraction, and the content absorption effect. Also, to collect qualitative data that may not be revealed by the questionnaire survey, semi-structured interviews were performed at the end of the digital curation exhibition activity. Additionally, the effects of the crowdsourcing and digital humanities with CAS support on collecting and organizing ideas and opinions for historical events and promoting humanity research outcomes were considered as our future works because they all need to take a long time to investigate.

Findings— Based on the questionnaire survey, this work found that the digital curation with CAS support revealed the highest rating score in terms of the item dimension of attraction effect. The result shows applying CAS to support digital curation is practicable, novel, and interesting to visitors. Additionally, this work also successfully applied the developed CAS to crowdsourcing and digital humanities so that the two research fields may be brought into a new ground.

Originality and value—Based on the CAS, this work developed a novel digital curation approach which has high satisfaction degree on attraction effect to visitors, an innovative crowdsourcing platform that combined with a digital archive system to efficiently gather collective intelligence to solve the difficult problems of identifying digital archive contents, and a high potential digital humanity research mode that can assist humanities scholars to annotate the texts with distinct interpretation and viewpoints on an ancient map as well as discuss with other humanities scholars to agitate more issues.

1. Introduction

In 2012, the Association of College and Research Libraries (ACRL) in the USA presented the development trend in academic libraries. As mentioned under the point “data curation”, standards for all types of data continue to evolve so that the data curation challenges are increasing. This leads to librarians and other information workers collaborating intimately with their research communities to facilitate the development of data curation. It is also mentioned that information technology continues to drive much of the futuristic thinking within academic libraries (ACRL, 2015). The issues mentioned above include the reflection on the challenge of a library’s role, which becomes more complicated, and the further evolvement with users under the digitalization trend. With the prosperous development of new technology, academic libraries must provide flexible, nimble, and responsive services for users, create a more efficient and convenient access to resources and information, and establish a new connection with academic resources. Consequently, academic libraries should provide more innovative academic resource services and integrate special collection materials, research resources, and digital information system with advanced technology.

Due to the influence of information digitalization and internet technology, the library curation techniques have gradually expanded from traditional curation of physical objects and collections into digital curation or the integration of substantial and digital curation. British Digital Curation Centre (DCC) proposed to deliberately take required resources into account before the display and suggested that “digital curation lifecycle” had to present definite work objectives at various stages, allow data being maintained, managed, and added value in the lifecycle, and present long-term preservation and accessibility (Digital Curation Centre). In recent years, Social Science Information Center (SSIC) of National Chengchi University Libraries in Taiwan, which is a modern research library, has already had great experiences and achieved remarkable success in digital curation. With the combination of information visualization techniques and the experiences in supporting the interaction with large display installations in several digital curations, SSIC presented the originally planar and static historical materials in a digital way. A prior study (Chen *et al.*, 2012) has successfully developed a reading annotation and knowledge sharing tool, which can annotate a web page with HTML format archived by the Taiwan libraries’ history digital library based on Web 2.0 technologies, to collect user-generated contents so that the digital library’s contents grow dynamically as readers contribute knowledge. In recent years, the developed reading annotation and knowledge sharing tool was further developed and enhanced as a collaborative annotation system (CAS) with more powerful annotating functions. In 2015, the collaborative annotation system (CAS) was applied to the digital curation in the Annual Forum of NCCU Digital Archives and Research 2015 for the first time, in order to encourage visitors to become actively involved in the interaction with the display. It enabled the visitors to supplement the content or express their thoughts so that the interaction between visitors and products could be enhanced, thus promoting an efficient knowledge transfer beyond the one-way information flow.

Furthermore, crowdsourcing is an emerging model for task dividing, problem solving and then goal achieving via the Internet (Brabham, 2008). It is also defined as the process of organizing by soliciting contributions from an undefined public, the organization or the company outsourcing the tasks to contributors from different backgrounds, acting on their own initiative (Whitla, 2009). In the fields of libraries, museums, and digital humanities, this kind of model also takes place in order to collect and organize ideas and opinions for historical events and social issues. To gather collective intelligence, SSIC also develops a crowdsourcing platform based on the CAS. Through the integration of the CAS with digital archives platform with digital resources, users can annotate and discuss the words and pictures in historical digital materials. Because of the collection of annotations, users are able to extract the common annotations within a certain range. With these common annotations, it will help the users engage in a more extensive research. It can be regarded as an interaction of content discovery, organization and sharing in order to approach to the goal of

information sharing and research facilitation. In addition, it is likely to help researchers restore the details of history and reconstruct the history by the common annotations of these amateurs. After extracting the common annotations from numerous users, research on the annotations in the databases and the records of discussion can also be made to investigate into the contributions of different users in annotation, discussion, and analysis of historical materials.

Research on digital humanities mostly interpreted historical materials via the support of digital technologies (Rushmeier, Pintus, Yang, Wong, and Li, 2015). Narrowly speaking, such a research model was called “digital computing” or “digital informatics” (Svensson, 2013). Nevertheless, the main value of research on humanities and social sciences still is on individual interpretation and analysis of humanities scholars because the ability of computers processing “meanings” is limited currently. Therefore, a more proper model should integrate various developed digital tools into the digital humanities research platform (DHRP) which covers long-term accumulated and stacked digital contents to provide humanities scholars with complete digital research environment as well as provides digital tools that can support data analysis and visualization techniques to assist humanities scholars’ research so that human-computer cooperation research model can be achieved (Sato, Goto, Kimura, and Maeda, 2016). Therefore, the SSIC of NCCU also devotes its efforts to integrate the developed CAS into DHRP to create custom functions so that researchers can sort, compile, interpret, and discuss digital materials conveniently. This function provided by the developed CAS will combine with interdisciplinary teams to expand the research orientations and to construct and share the knowledge towards the achievement of sharing academic outcomes. In addition to knowledge share and the enhancement of research convenience, it is expected to allow humanities scholars agitating innovative problem awareness and restoring a brand-new text context through annotation consensus to discover more history details in historical archives and assist humanities scholars in reconstructing history (Renzel, Cao, Lottko, and Klamma, 2010). After extracting the annotated interpretation of several humanities scholars, it is expected to, aiming at the dialectical data of professional users, further construct and share knowledge by expanding distinct research dimension with data mining.

In summary, this work presents a novel CAS and discusses its applications in digital curation, crowdsourcing, and digital humanities because of existing strong relations among them. Michael (2016) indicated that the convergence of new technologies in digital humanities and the growing importance of open access in the research arena lead to an increasing focus on how archival and special collections are made available digitally, are curated, and are preserved. Moreover, Dunn and Hedges (2012) identified four factors that define crowdsourcing used within digital humanities research. The first factor is a clearly defined core research question and direction within the digital humanities; the second factor is the potential for an online group to add, transform, or interpret data that is important to the digital humanities; the third factor is a definable task which is broken down into an achievable workflow; and the fourth factor is the setting up of a scalable activity which can be undertaken with different levels of participation. Obviously, digital curation can provide well-organized and added value digital contents to humanity scholars, thus being regarded as a basis of the research of digital humanities, whereas crowdsourcing is a useful method to gather collective intelligence for the research target of digital humanities.

2. Background

Annotations are a form of active reading behavior and reflect readers’ thinking processes. The act of annotation enhances memorization and enables learners to think about and solve problems (Marshall, 1997). Annotations facilitate students’ learning and cognitive processes and enable them to convert their basic prior knowledge into high-level cognitive patterns (Kiewra, 1989). Marshall (1997) addressed the six major purposes of annotation, which are to emphasize, to highlight a key point, to improve understanding, to explain, to notify of a problem to be solved, and doodling.

Additionally, annotations help readers to memorize, analyze, and clarify contents; of these functions, memorizing most benefits readers (Ovsiannikov, Arbib, and Mcneill, 1999).

In recent years, collaborative annotation systems, such as Diigo (Estellés, Moral, and González, 2010), PAMS (Su, Yang, Hwang, and Zhang, 2010), and CRAS-RAIDS (Chen and Chen, 2014), allows learners to collaborate efficiently in annotating digital texts in order to add valued information, share ideas by expressing different perspectives on digital texts with annotations, as well as create knowledge by reading digital texts with annotations. These collaborative annotation systems have been successfully applied in promoting learners' collaborative reading performance (Chen and Chen, 2014; Su, Yang, Hwang, and Zhang, 2010). Furthermore, Sato, Goto, Kimura, and Maeda (2016) developed a prototype Web-based system which has features in terms of supporting multiple users to make annotations to the same document simultaneously and suggesting annotation function based on existing annotation strings for collaboratively making annotations on historical documents by multiple humanities researchers who are distant from each other. However, few collaborative annotation systems provide simultaneously four types of multimedia annotations including text annotation, picture annotation, voice annotation, and video annotation to support collaborative annotation. Also, most of collaborative reading annotation systems lack reading annotation and interactive discussion scaffolds that can improve reading performance in collaborative digital reading environments. Therefore, this work proposes a novel CAS which simultaneously provides four types of multimedia annotations, the reading annotation scaffold, and the interactive discussion scaffold to improve collaborative annotation. More importantly, although collaborative annotation system has been preliminarily applied in digital humanities, how to successfully develop a practicable mode of using collaborative annotation system to increase the academic productions of digital humanity researchers is still a valuable research issue. Finally, to the best of our knowledge, the study of applying collaborative annotation system in crowdsourcing and digital curation still lacked. The study aims to fill the above research gap.

3. Methodology

3.1 The Implemented CAS System

3.1.1 The System Architecture of the Proposed CAS

Refer to the studies of Wolfe (2002) and Chen *et al*. (2012), this work proposed a novel architecture of CAS as shown in Figure 1. This work adopted Asynchronous JavaScript (Ajax) and Model-View-Controller (MVC) framework to implement a CAS with the reading annotation tools for knowledge creating, archiving, and sharing services so that readers can perform collaborative reading annotations for any HTML web pages. The system architecture is composed of three parts including the Browser Client, Data Host, and CAS Server. The Browser Client is responsible for providing a friendly user interface to service a reader who visits the Data Host for annotating a HTML web page. The Data Host is in charge of providing the static HTML web page for readers to read and annotate. The CAS Server aims at providing collaborative reading annotation functions. The Data Host must first install the CAS display element (i.e. View) by using Ajax syntax to enable the reader to browse and annotate HTML web pages provided from the Data Host. The CAS display components installed on the Data Host must regularly communicate with the CAS server, while the Controller and Model components in the CAS Server play the role of storing and providing annotation data. Based on the system architecture, the system operating procedure is described and summarized as follows.

Step 1. Login system: A reader logins the CAS via the Browser Client to annotate a HTML web page provided from the Data Host.

Step 2. Request reading data: The display components of the Data Host send a request of processing a HTML web page to the Controller of the CAS Server to get reading data.

Step 3. Process reading data: The Controller of the CAS Server processes the reading data coming from the Model components of the CAS Server.
Step 4. Respond the processed data: The Controller of the CAS Server responds the processed data to the display element of the Data Host.
Step 5. Display annotation: The display element of the Data Host shows the annotation results contributed from the reader on the user interface of the Browser Client.

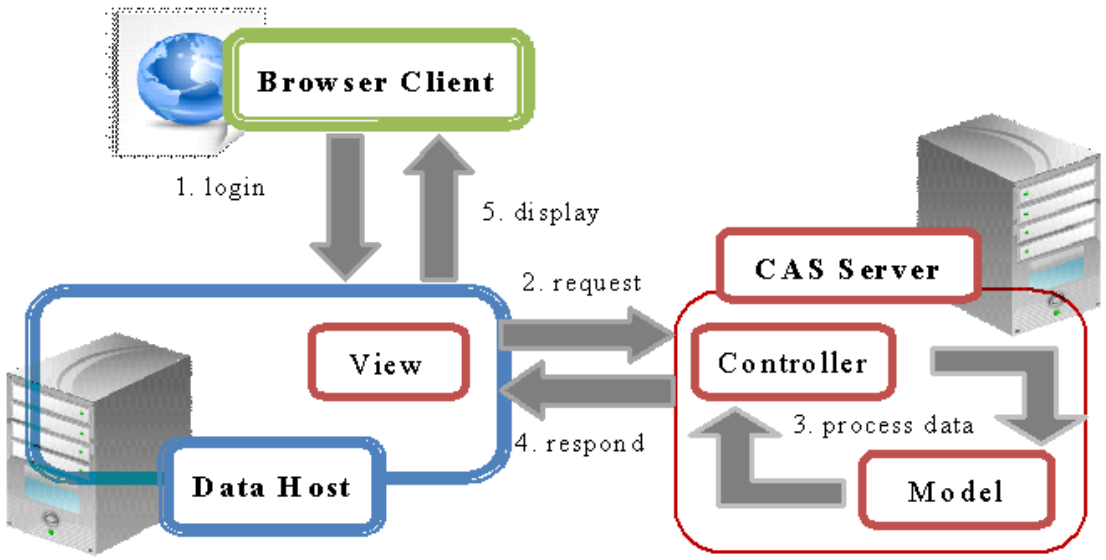

Figure 1. The System Architecture of the Proposed CAS

3.1.2 The System Functions of the Proposed CAS

The proposed CAS provides four main functions to support reading annotations, as explained in the following:

(1) Create/Modify/Delete the reading annotation
The proposed CAS provides a reading annotation tool with a friendly graphical user interface that can help readers create/modify/delete annotations for a HTML web page. To help readers construct useful annotations, the CAS provides four types of multimedia annotations including text annotation, picture annotation, voice annotation, and video annotation (Fig. 2). Using text type annotation, learners can directly type additional text information in the text window for the selected text of the reading HTML web page. Additionally, the CAS system provides reading annotation scaffolds that can be predetermined arbitrarily by administrators for the text type annotation (Fig. 2). For example, the reading annotation scaffolds of the text type annotation can be determined as discussion, question, place, and figure for a digital humanity application scenario. Using picture type annotation, learners can select a rectangle area on a picture to add additional text information. Using voice type annotation, learners can record their spoken words for the selected text of the reading HTML web page via a microphone. Using video type annotation, learners can link an URL address of a selected video associated with the selected text of the reading HTML web page from Youtube to provide additional information.

(2) Mark the favorite reading annotation
The proposed CAS allows readers to mark their favorite annotations. This feature helps CAS identify popular annotations when reviewing annotations of individual readers.

(3) Set the reading annotation mode
The proposed CAS has two reading annotation modes: the private reading annotation mode that

the reading annotations created by some readers can only be browsed by the annotation creator and the public reading annotation mode that the reading annotations created by some reader can be browsed by all readers.

(4) Discuss the reading annotation

The proposed CAS provides a corresponding discussion board for each public reading annotation (Fig. 3). Restated, a reader can respond to a specific reading annotation contributed by other readers by posting texts to the discussion board in order to debate, discriminate, or communicate ideas associated with the reading annotation.

Figure 2. The user interface of the CAS with four types of multimedia annotations including text annotation, picture annotation, voice annotation, and video annotation

Figure 3. The user interface of the discussion board for public reading annotations

3.2 Research Participants

In total, 45 visitors who participated in the digital curation with CAS support in the exhibition of the forum entitled “Tides of Humanity and Literature in Post-War Taiwan” were regarded as research participants. Since conducting questionnaire survey and interview for all the visitors who took part in the digital curation with CAS support is extremely difficult, a total of 15 visitors were randomly invited to participate in questionnaire survey and semi-structured interviews. Among 45 research participants, 20 (45%) were male and 25 (55%) were female. Participants were aged 23–56. All research participants had no experiences on the digital curation with CAS support, but most of them had experiences on the digital duration with the support of information visualization technologies and large interactive display devices.

3.3 Research Instrument

A visitor satisfaction questionnaire was designed to assess three dimensions of user satisfaction for the digital curation with CAS support. The three dimensions were the interactivity with displayed products, the attraction, and the content absorption effect. The visitors responded to each questionnaire item using a 5-point Likert-type scale ranging from 1 for “strongly disagree” to 5 for “strongly agree.”

3.4 Research Design

According to referring the studies of Wolfe (2002) and Chen *et al*. (2012) associated with developing reading annotation system, this study proposed a novel CAS that can help readers perform collaborative reading annotations for any HTML web pages. Moreover, this study applied successfully CAS in the applications of digital curation, crowdsourcing, and digital humanities. To confirm the satisfactory of visitors who experienced the digital curation with CAS support, the

forum entitled “Tides of Humanity and Literature in Post-War Taiwan” first applied the CAS technology to combine with e-books to present novel digital curation style on digital display devices in 2015. After all the visitors experienced the digital curation with CAS support, a questionnaire survey method were used to investigate the ideas and satisfaction of visitors who participated in the digital curation with CAS support in the item dimensions of the interactivity with displayed products, the attraction, and the content absorption effect. Additionally, to collect qualitative data that may not be revealed by the questionnaire survey, semi-structured interviews were performed at the end of the digital curation exhibition activity. Exploiting the inherent flexibility of a semi-structured interview, the interviewer reused or repurposed questions to obtain in-depth information on the perspectives and personal experiences of each interviewee. Additionally, the evaluation of the crowdsourcing and digital humanities with CAS support was considered as our future works because they all need to take a long time to investigate. Finally, this study summarizes the research results and proposes several potential future research directions.

4. Application of CAS to Digital Curation

Based on the promotion of digital archives and the expansion of research and development outcomes on humanities and social science, the SSIC of NCCU has combined information visualization technologies and matched with large interactive display devices to develop digital curation technologies so that the originally planar and static historical data can be demonstrated by digital display technology. Besides, the forum entitled “Tides of Humanity and Literature in Post-War Taiwan” first applied the CAS technology to combine with e-books to present a novel digital curation type on digital display devices in 2015. The following five types of digital curation for the forum were produced.

4.1 Discussion of Cubism-CAS Display

The first type of digital curation produced for the forum is based on integrating digital contents with the CAS. Collaboratively annotating digital texts allows users to add valued information, share ideas, and create knowledge. The CAS has functionalities including editing, revising, deleting, browsing, sharing annotations, and on-line discussion for digital texts with annotations. Selecting Kuo-song Liu’s article, “Discussion of Cubism”, published in Bihui magazine as the example, a user could add annotations for any selected texts to provide and share extra information via the CAS (Fig. 4). The types of annotations can be text, voice, image, and video. Figure 5 shows an example of the type of text annotation on “Discussion of Cubism”.

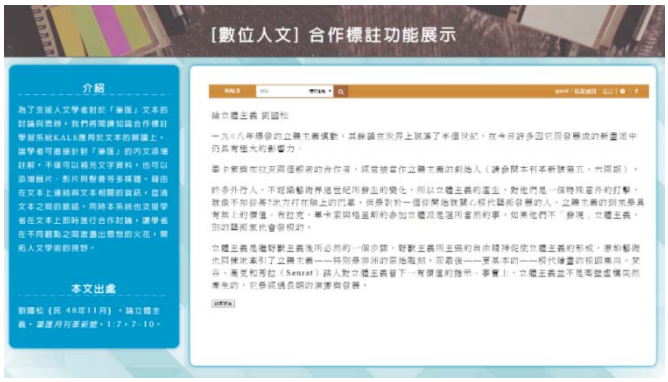

Figure 4. Screenshot of “Discussion of Cubism” on the CAS display

Figure 5. An example of the type of text annotation on “Discussion of Cubism”

4.2 Alternation Display of Related Photos of Professor Yu on Digital Photo Wall

The second type of digital duration produced for the forum is based on digital wall with alternation display function. The digital photo wall integrates digital photo carousel and solid calligraphy. Based on the digital wall with alternation display, the digital display of “Related photos of Professor Yu” is produced (Fig. 6).

Figure 6. Alternation display of related photos of Professor Yu on digital wall

4.3 Interaction with the Alternation of Photos of Professor Yu and His Friends and Magazine Covers on a Wide Projection Wall

The third type of digital curation produced for the forum is based on a wide projection wall composed of several projectors, on which high-resolution photos are combined with multi-point touch modules allowing several visitors freely browsing simultaneously and further interacting with the image content (Fig. 7).

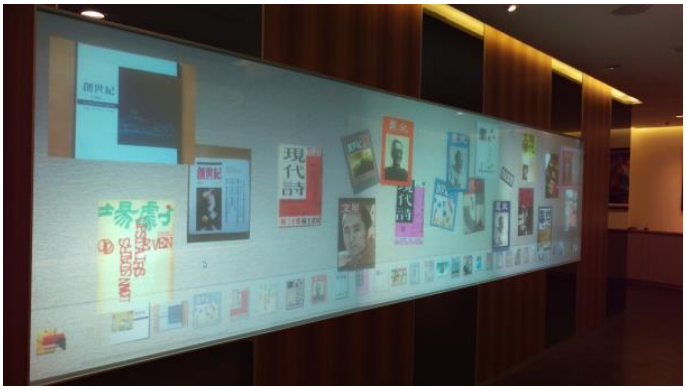

Figure 7. Interaction with the alternation of photos of Professor Yu and his friends and magazine covers on a wide projection wall

4.4 Trailer of the Inspired Island: Series of Eminent Writers from Taiwan on an Interactive Film Display Interface

The fourth type of digital curation produced for the forum is based on an interactive film display interface. By integrating various design resources, this type of digital demonstration plays the interviews with writers in The Inspired Island: Series of Eminent Writers from Taiwan, authorized by Fisfisa Media (Fig. 8).

Figure 8. Trailer of The Inspired Island: Series of Eminent Writers from Taiwan on an interactive film display interface

4.5 E-books: Introduction of Bihui Magazine and Important Authors

The fifth type of digital curation produced for the forum is based on E-books. Focusing on Bihui, Literature Quarterly, Literature Bimonthly, and Literature published by Professor Tian-Cong Yu and the important authors (Fig. 9), two types of e-books were designed to make visitors deeply understand and experience such journals.

Figure 9. Introduction of Bihui and important authors on E-books

In sum, this forum expected to break the one-way information demonstration channel, manifest the characteristic of interaction between visitors and displayed products, and promote the communication efficacy of richer knowledge by offering the innovative academic resource services.

4.6 Assessing Satisfactory of Visitors Who Experienced Five Types of Digital Curation

A total of 15 visitors who participated in the forum were randomly invited to fill out the questionnaire, which aims to understand the ideas about the five types of digital curation in this forum by investigating the interactivity with displayed products, the attraction, and the content absorption effect. Moreover, this work also designed an open-ended question to allow visitors expressing their comprehensive opinions. Table 1 shows the descriptive statistics results of visitor satisfaction for five types of digital curarion. Analytical results show that “Alternation display of related photos of professor Yu on digital wall” was ranked on the top in terms of interactive effect among the five types of digital curation produced for the forum, followed by “Trailer of The Inspired Island Series of Eminent Writers from Taiwan on an interactive film display interface”, and then “E-books: Introduction of Bihui Magazine and Important Authors”. “Study on cubism-CAS display” was ranked on the top in terms of attraction effect, followed by “Alternation display of related photos of professor Yu on digital wall”, and then “Interaction with the alternation of photos of Professor Yuand his friends and journal covers on a wide projection wall”. “Alternation Display of Related Photos of Professor Yu on Digital Photo Wall” was ranked on the top in terms of content absorption effect, followed by “Trailer of the inspired island: Series of eminent writers from Taiwan on a wide projection wall”, and then “Interaction with the alternation of photos of professor Yu and his friends and magazine covers on a wide projection wall”. “Alternation display of related photos of professor Yu on digital wall” was ranked on the top in terms of total effect among the five types of digital curation produced for the forum, followed by “Study on cubism-CAS display”, and then “Trailer of the Inspired Island: Series of Eminent Writers from Taiwan on an Interactive Film Display Interface”.

15 visitors who were interviewed gave a lot of positive opinions and feedback on the type of digital curation with the CAS support produced for this forum. In comparison with traditional seminars, which emphasize more on the exchange and interaction among authors, commentators, and audience, most of them considered that digital content display with the CAS support could further provide participants and researchers with possibilities to interact with “concretized” hard copies and historical materials. Visitors no longer passively listen to lectures but could more deeply experience the theme of the forum and largely reduce the difficulty in reading and data collection by the lecturers reporting the academic papers matched with periphery digital content display to show the presence. It could also promote visitors’ interests to offer better ideas and orientations for future research. Since applying the CAS display to digital curation is the first time, interviewees also proposed some improvement suggestions. Interviewees identified the rich annotations and acquired wonderful information from the CAS, but mentioned the small screen window, whose size could not be adjusted, the obstacle to clicking on annotations, and the adverse reading, and suggested designing the annotation window on the left column for convenient reading of users at various ages.

Table 1. Descriptive statistics results of visitor satisfaction for five types of digital curarion

Type of Digital Curation	Dimension	Number of Visitors	Mean	Std.
Discussion of Cubism-CAS Display	Interactivity	15	2.47	.990
	Attraction	15	**4.47**	**1.060**
	Absorption	15	2.87	1.407
	Total	15	9.80	2.484
Alternation Display of Related Photos of Professor Yu on Digital Photo Wall	Interactivity	15	**3.80**	**1.424**
	Attraction	15	3.53	1.552
	Absorption	15	**3.60**	**1.682**
	Total	15	**10.93**	**4.559**
Interaction with the Alternation of Photos of Professor Yu and His Friends and Magazine Covers on a Wide Projection Wall	Interactivity	15	2.53	1.598
	Attraction	15	2.67	1.676
	Absorption	15	2.93	1.100
	Total	15	8.13	4.224
Trailer of the Inspired Island: Series of Eminent Writers from Taiwan on an Interactive Film Display Interface	Interactivity	15	3.13	1.356
	Attraction	15	2.47	1.187
	Absorption	15	3.20	1.424
	Total	15	8.80	3.877
E-books: Introduction of Bihui Magazine and Important Authors	Interactivity	15	2.67	1.234
	Attraction	15	2.53	1.302
	Absorption	15	2.13	1.187
	Total	15	7.33	3.155

5. Application of CAS to Developing Crowdsourcing Platform

5.1 Introduction of Crowdsourcing Platform with Annotation Function

The crowdsourcing platform is based on NCCU Memory website (Fig. 10), which is the archives of NCCU members' common memory, allowing the faculty, staff and student of National Chengchi University to establish albums and freely upload images and photos related to National Chengchi University. In addition to the images and photos of activity in past years and campus corners, plenty of university historical photos are also contained. Such photos archive not only campus landscape, campus activity, and student life, but also records of important visitors. Users could randomly click on the albums in the photo area, select specific photos through the popular tag auto-screening offered by the website, or search for photos with keywords.

Figure 10. NCCU Memory homepage

The CAS was also mounted on the NCCU Memory website. From the red square shown in Fig. 11, the CAS toolbar was added on the top of the NCCU Memory website, where a user could use the CAS functions with simply registration.

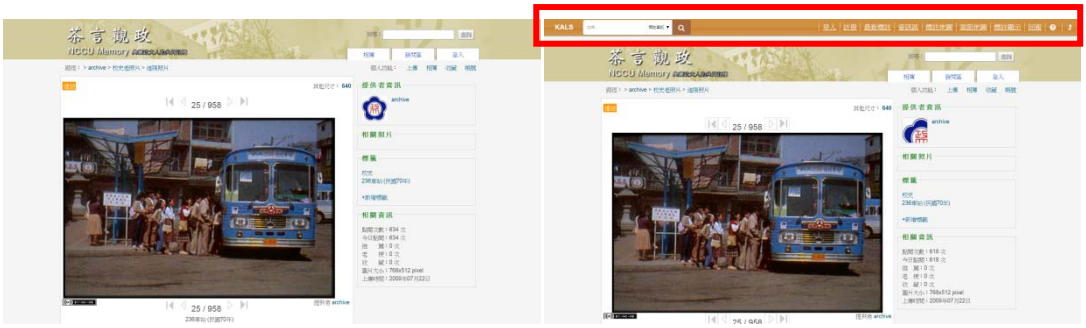

Figure 11. Original frame (left) of NCCU Memory photo browsing webpage and mounted CAS (right) (red square shows the CAS)

A user, who would like to annotate photos, could move the mouse cursor to the photo, on which an Add Note button appears (Fig. 12). The user can click on the button, and edit annotations on the annotation window (Fig. 13). After annotating the photo, the user simply clicks on the annotation block to browse the annotation content (Fig. 14). Such a function allows users expressing opinions, proposing questions, or sharing known information (Fig. 15), and further exploring the changes of spatio-temporal background, figures, or new and old buildings in a photo.

Figure 12. Moving the mouse cursor to the photo on which an Add Note button appears

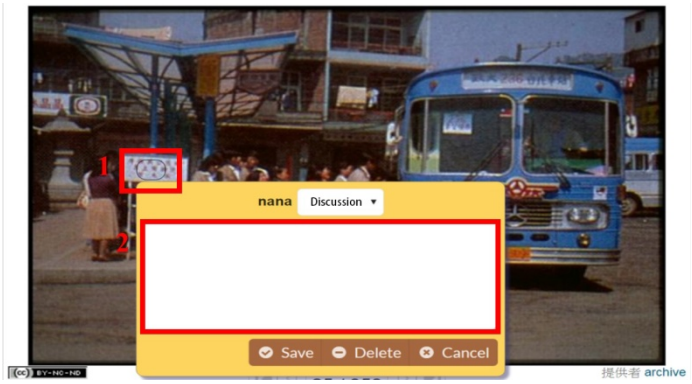

Figure 13. Photo annotation editing window (1) annotation block: the annotation location and annotation block size could be adjusted; (2) annotation contents key in

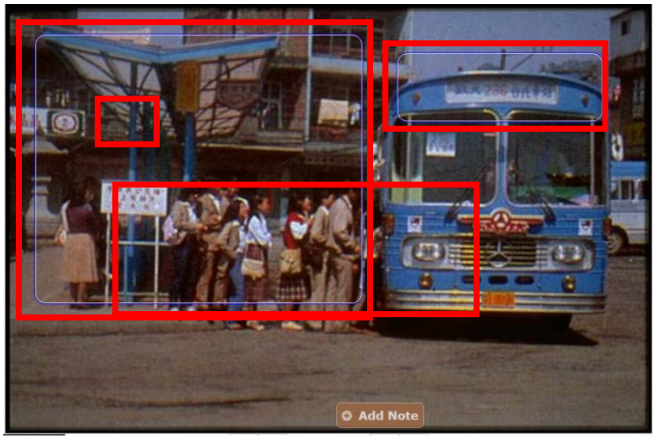

Figure 14. Moving the mouse cursor to the photo on which all annotated blocks are displayed

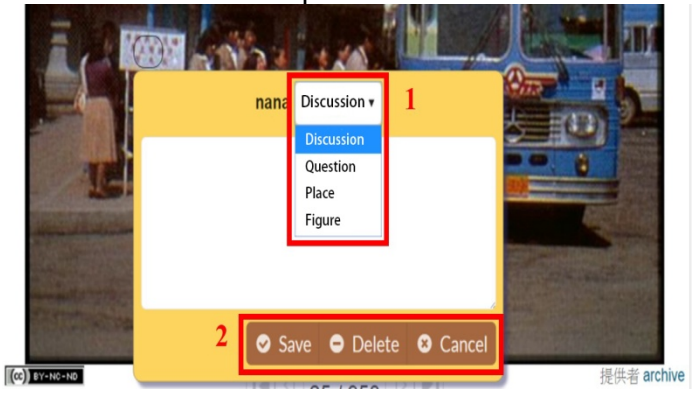

Figure 15. Photo annotation editing window (1) pull-down menu for selecting the four types of

annotations including discussion, question, place, and figure; (2) options of save/delete/cancel annotations

5.2 Application and Future Development of Crowdsourcing Platform with Annotation Function

During the system testing stage, librarians of National Chengchi University were invited to use the NCCU Memory website to perform crowdsourcing with the CAS support in order to collect their knowledge and memory. In this case, not only do the libraries appear resonance with the university history, but the precious historical photos are also added rich descriptions so that the stories of pure photos are expanded through crowd power.

In order to extract users' knowledge, it is expected to acquire users' annotation records, process with statistical computation, extract crowd consensus from massive annotation contents in the future. Moreover, it is expected to expand the user range by combining the activity of library theme week and anniversary of the university and applying community media resources, and increasing reward and feedback systems to promote the participation of crowd.

6. Application of CAS System to Digital Humanities

6.1 Introduction of CAS-based Digital Humanities Research Platform Function

Taking the ancient map of Tamsui in Taiwan as the sample, the image acquired from Research Center for Humanities and Social Sciences of Academia Sinica – Map of hundred years of Tamsui history (http://gissrv4.sinica.edu.tw/gis/tamsui.aspx) was input to the CAS-based digital humanities research platform. A humanities scholar could annotate the historical texts with distinct interpretation and viewpoints through the CAS functions on the digital humanities platform as well as discuss with other humanities scholars to agitate more issues (Fig. 16).

Figure 16. Homepage of CAS-based digital humanities research platform

6.1.1 Photo and text annotation function

This platform offers photo annotation functions. As the example of the ancient map of Tamsui in Japanese ruled period, a user could freely annotate the name of place in Tamsui during Japanese ruled period on the map after logging in the platform. When using the scanned photo on the historical text, a humanities scholar could directly annotate the interpretation and opinions about the

article content on the photo and discuss with other scholars (Fig. 17).

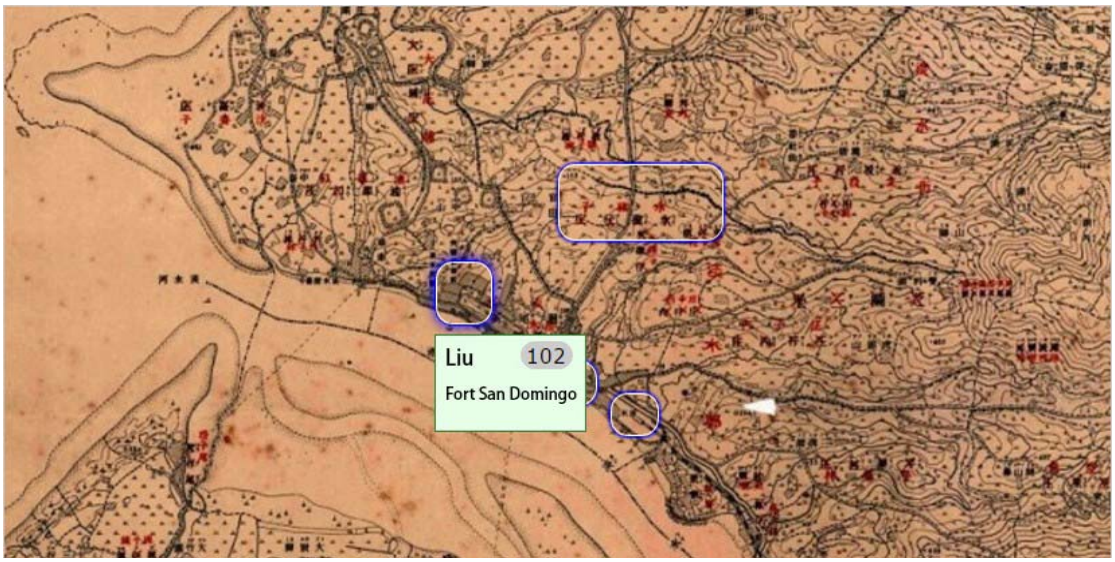

Figure 17. Photo annotation function of CAS-based digital humanities research platform

Furthermore, a user could make text annotations after logging in the digital humanities research platform (Fig. 18). The user could select any length of words on the article to start the annotation, and the types of annotations could be adjusted according to the research requirement. In the example in the Fig. 18, the types of annotations include discussion (red), question (orange), place (yellow), and figure (green). “Discussion” means to discuss the article; “question” is to annotate some confusing vocabulary and expect other users to solve the problem; “place” refers to explaining places appeared on the ancient map or articles with annotations; and, “figure” aims to describe people in photos or articles with annotations.

Figure 18. Text annotation function of CAS-based digital humanities research platform

6.1.2 Annotation map and section map

This platform provides annotation map functions, allowing presenting the quantity and distribution of various types of annotations on the annotation map and viewing the discussion hot spots on the article content by the number of annotations. Clicking the type of annotation on the

annotation map allows rapidly moving to the location with corresponding annotation as well as clicking the paragraph name on the section map allows rapidly moving the corresponding location of the paragraph for reading the text contents (Fig. 19).

Figure 19. Annotation paragraph distribution of CAS-based digital humanities research platform

A humanities scholar could overturn the past research model of independently interpreting and analyzing historical texts through the platform and agitate more new research questions by discussing with other scholars (Fig. 20).

Figure 20. Annotation discussion of CAS based digital humanities research platform

6.2 Future Development and Application of Digital Humanities Research Environment Platform

To offer favorably collaborative research environment for humanities scholars and further explore the complicated interpersonal relationship and social context hidden in the texts interpreted by humanities scholars, the CAS-based digital humanities research platform could assist researchers in discussing, interacting, and interpreting historical texts and could have the following deeper applications.

(1) Data Mining

The SSIC of NCCU expects to acquire the annotation texts and annotation records accumulated

by humanities scholars on the CAS-based digital humanities research platform. Through data mining technologies, the text mining and machine learning techniques can be used to discover the useful knowledge from the large amount of annotations on the historical texts.

(2) Social network analysis

In the research on humanities and social sciences, independent analyses and interpretation of humanities scholars are the most important, in addition to the historical value of historical materials. Researchers could interpret historical texts on the platform as well as respond to comment on other scholars' points of view. As a result, a lot of interactive data would be generated in the collaborative process. Accordingly, the SSIC of NCCU expects to apply the above data to analyze the social cooperation networks among scholars and facilitate the cooperation among more relevant researchers, or attempt to analyze the knowledge characteristics with scholars' professional specialties.

(3) Process record analysis

Recording the operation processes of humanities scholars on the platform to evaluate the effectiveness and ease of use of the platform could improve and reinforce the platform functions, analyze users' behaviors, and expect to find out the optimal interactive and collaborative research model for humanities scholars in the future.

7. Conclusions and Future Work

This paper presents a novel CAS and successfully applies it to digital curation, crowdsourcing, and digital humanities. In the future, it is expected to optimize the CAS, integrate a digital archive system such as DSpace, and develop different digital tools constantly, such as text mining, machine learning, and statistical data analysis tools, to develop the digital humanity research platform (DHRP) in order to lead the researchers to undertake the digital quantitative research and create new knowledge. Besides, it is also expected to gather as many academic research resources as possible in the university and turn them into the DHRP so that DHRP gradually becomes as an important humanities and social study resource platform.

The SSIC of NCCU also expects to promote the librarians cooperate with various users in the research and development process of applying CAS to digital curation, crowdsourcing, and digital humanities studies so that the librarians see the demands of users and inspect the opinions and feedback of users, the users become the knowledge promoters and consultants, and the users with distinct experiences and background create exchange opportunities. Accordingly, the development experience and the future prospect are concluded as follows:

(1) Continuously reinforce the cooperation with academic units inside and outside the university

In the future, when the system functions of CAS are developed more completely, the SSCI of NCCU will build partnership with academic units, encourage users to use the CAS, develop the multiple effects of knowledge communication, and link and integrate the professional resources to generate more innovative outcomes.

(2) Optimize the CAS function

When users proceed the collaborative reading annotation, disperse annotation information which is more complicated and hard to be comprehended is likely generated. For this reason, it is expected to design the reading navigation map in the CAS for helping users master the distribution locations and quantity of annotation content and assisting users in generating richer perception and deeper acquaintance. It also manifests the value-added function of the text data generated by users.

Furthermore, it is expected to develop an annotation recommendation system that can automatically recommend the annotation contents that users might be interested to enhance personalized services.

(3) Integrate crowdsourcing platform with community media to enhance communication and cooperation

Integrating CAS-based crowdsourcing platform with community media, such as Facebook, Twitter, and Instagram, allows users sharing annotated web pages with their friends or tracers to promote the exposure of crowdsourcing content and attract potential users. This development will benefit the communication and cooperation between the library and users through users' recommendation.

(4) Continuously develop various digital analysis functions in the DHRP

Undoubtedly, the DHRP can guide researchers to create new knowledge, facilitate teamwork among relevant researchers. In the future, the SSIC of NCCU will also devote to developing digital analysis technologies, such as data mining, social networks analysis, and process record analysis, in order to offer humanities researchers with complete digital content and technology integrated DHRP so that the DHRP becomes a critical digital humanity research platform around the world.

References

ACRL (2015), "Environmental Scan 2015", available at: http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/EnvironmentalScan15.pdf (accessed 4 May 2016).

Brabham, D. C. (2008), "Crowdsourcing as a model for problem solving an introduction and cases", *Convergence: The International Journal of Research Into New Media Technologies*, Vol. 14 No.1 , pp. 75-90.

Chen, C. M., Chen, Y. T., Hong, C. M., Liao, C. W., and Huan, C. M. (2012), "Developing a Taiwan Library History Digital Library with Reader Knowledge Archiving and Sharing Mechanisms Based on the DSpace Platform", *The Electronic Library*, Vol. 30 No. 3, pp. 426- 442.

Chen, C. M. and Chen, F. Y. (2014), "Enhancing digital reading performance with a collaborative reading annotation system", *Computers & Education*, Vol. 77, pp. 67–81.

Digital Curation Centre, "What is digital curation", available at: http://www.dcc.ac.uk/digital-curation/what-digital-curation (accessed 4 May 2016).

Dunn, S. and Hedges, M. (2012), "*Crowd-sourcing scoping study. Engaging the crowd with humanities research"*, Report for the UK Arts and Humanities Research Council Connected Communities Scheme, available at: http://crowds.cerch.kcl.ac.uk/wp-content/uploads/2012/12/Crowdsourcing-connected-communities.pdf (accessed 4 May 2016).

Estellés, E., Moral, E., and González, F. (2010), "Social bookmarking tools as facilitators of learning and research collaborative processes: the Diigo case", *Interdisciplinary Journal of E-Learning and Learning Objects*, Vol. 6, pp. 175-191.

Kiewra, K. A. (1989), "A review of note-taking: The encoding-storage paradigm and beyond", *Educational Psychology Review,* vol. 1, no. 2, pp. 147-172.

Marshall, C. C. (1997), "Annotation: From paper books to the digital library", In Proceedings of *the Second ACM Conference on Digital Libraries*, pp. 131-140.

Michael, J. P. (2016), "Digital curation in the digital humanities: Preserving and promoting archival and special collections", *The Electronic Library*, Vol. 34 No. 6, pp. 1055-1055.

Ovsiannikov, A. I., Arbib, M. A., and Mcneill, T. H. (1999), "Annotation Technology. *International Journal of Human-Computer Studies"*, Vol. 50, pp. 329-362.

Renzel, D., Cao, Y., Lottko, M., and Klamma, R. (2010), "Collaborative video annotation for

multimedia sharing between experts and amateurs", In A. Carreras, J. Delgado, X. Maronas, V. Rodrı`guez (Eds.), *Proceedings of the 11th international workshop of the multimedia metadata community on interoperable social multimedia applications (WISMA-2010)*. vol. CEUR workshop proceedings (pp. 7–14). Barcelona, Spain: CEUR Workshop Proceedings.

Rushmeier, H., Pintus, R., Yang, Y., Wong, C., and Li, D. (2015), "Examples of challenges and opportunities in visual analysis in the digital humanities", available at: http://graphics.cs.yale.edu/site/sites/files/dh-ei-SUBMIT.pdf (accessed 4 May 2016).

Sato, T., Goto, M., Kimura, F., and Maeda, A. (2016), "Developing a collaborative annotation system for historical documents by multiple humanities researchers", *International Journal of Computer Theory and Engineering*, Vol. 8 No. 1, pp. 88-93.

Svensson, P. (2013), "Humanities computing as digital humanities", In *Defining Digital Humanities: A Reader*, pp. 159-186.

Su, A. Y. S., Yang, S. J. H., Hwang, W. Y., and Zhang, J. (2010), "A Web 2.0-based collaborative annotation system for enhancing knowledge sharing in collaborative learning environments", *Computers & Education*, Vol. 55 No. 2, pp. 752-766.

Whitla, P. (2009), "Crowdsourcing and its application in marketing activities", *Contemporary Management Research*, Vol. 5 No. 1, pp. 15-28.

Wolfe, J. (2002), "Annotation technologies: A software and research review", *Computers and Composition*, Vol. 19 No. 4, pp. 471-497.