Review article

Current status of digital humanities research in Taiwan

Pan Yuting [a], Jiang Yinfeng [b],[**], Zhang Jingli [b],[*]

[a] Nankai University, Department of Information Resources Management, Business School, Tianjin 300071, PR China
[b] Army Medical University Library, Chongqing 400038, PR China

A R T I C L E   I N F O

A B S T R A C T

Purpose: Review the current research status of the theory, techniques, and practice of digital humanities in Taiwan.
Methods: Select the 8 issues of the Journal of Digital Archives and Digital Humanities from its inception in 2018–2021, and the papers of the 5-year International Conference of Digital Archives and Digital Humanities from 2017 to 2021 as the research data, and conduct text analysis of the collected 252 articles.
Results: From the statistical analysis results, the number of practical articles is the largest, followed by tools and techniques, and the least number of theoretical articles. Text tools and literature research are the most concentrated aspects of digital humanities research in Taiwan.
Limitations: It still needs to be further compared with the current research status of digital humanities in Mainland China.
Conclusions: Digital humanities in Taiwan focuses on the development of tools and techniques, and practical applications of literature and history, and focuses on Taiwan's native culture to form its own digital humanities research characteristics.

1. Introduction

With the rapid development of information technology, digitization has increasingly become an important means of document preservation, dissemination, and utilization. Under this trend, Taiwan has carried out a digital collection plan since 2002. The digital collection program aims to use digital technology to collect valuable documents, promote the diversified application of documents, and promote the dissemination of culture. Since the implementation of the plan, Taiwan University, Academia Sinica, and other institutions have successively set up digital collection centers, which have produced rich digital achievements and established more than 100 kinds of databases. Some typical databases and access addresses are listed below.

1. Taiwan History Digital Library, http://thdl.ntu.edu.tw/ [1].
2. National Palace Museum Digital Library of Qing Arvhives, https://qingarchives.npm.edu.tw/ [2].
3. Context Discovery System for LIDAIBAO'AN, http://lidaibaoan.digital.ntu.edu.tw/ [3].

Digital technology has realized the preservation, dissemination and utilization of precious documents and materials, which has led

---

* Corresponding author.
** Corresponding author.
  E-mail addresses: pyt@mail.nankai.edu.cn (P. Yuting), jlzhang@tmmu.edu.cn (Z. Jingli).

to the discussion on the deep development and use of digital archives: how to further use digital archives and materials in education, society, academic research and other aspects? How to expand the educational function and social significance of precious documents and materials? And how to promote the deep-seated promotion and dissemination of culture? With the introduction of the concept of digital humanities into academic circles in Taiwan and the continuous discussion of the above issues, various research institutions have started digital humanities (DH)research one after another. The main sign of the wave of digital humanities research in Taiwan is the holding of the first *International Conference of Digital Archives and Digital Humanities* in 2009 [4]. The International Conference of Digital Archives and Digital Humanities (DADH) was initiated by the Research Center for Digital Humanities of National Taiwan University in 2009. Up to now, it has held 12 annual academic conferences, and it is also the earliest international conference on digital humanities held in Asia, which is of great significance to the development of digital humanities in recent years. The 13th DADH will be held on December 10–12, 2022 at the College of Arts, National Chung Hsing University.

By observing the objectives of previous DADH conferences, we can find that DADH2009 to DADH2012 can be regarded as the initial exploration period of DH study in Taiwan. While continuously accumulating abundant digitized resources, they are considering how large-scale digital data can facilitate knowledge creation [5–7]. Until DADH 2014 [8], the conference object at *Crossover & Transformation*, hoping to involve more researchers to participate in and promote the integration of expertise from different disciplines, thus inspiring more new research topics. DADH2014 can be seen as a symbol that digital humanities research in Taiwan has realized the transformation from digital collection to digital humanities. DADH2014 can be seen as a symbol that DH research in Taiwan has realized the transformation from digital collection to digital humanities. DADH2015 [9] and DADH2016 [10] basically continue the theme of DADH2014, indicating that the 2014 seminar has laid the scope and vision of DH research since then, and 2014–2016 can be regarded as the preliminary accumulation stage of DH research in Taiwan. In 2017, the 8th DADH took *Digital Humanities Evolving: Past, Present, and Future* as its theme [11], marking a new stage of DH research in Taiwan.

These *International Conference of Digital Archives and Digital Humanities* are the most important positions of digital humanities research in Taiwan. The academic activities taking the conference as an opportunity have produced many theoretical achievements, tools and techniques, and practical cases. By fully studying these contents, we can understand the current research status of digital humanities in Taiwan. Taking these into account, we decide to start from DADH2017 to conduct in-depth and comprehensive research on DH activities in Taiwan as far as possible, and fully demonstrate the development and characteristics of DH research in a specific region, so as to provide an important reference for other regions or academic communities to carry out DH research in the future.

2. Literature review

The combination of modern computer and network technologies with the traditional humanities research has given birth to a new research field called Digital humanities (DH) [12]. The emergence of DH not only promoted the change of research paradigm, but also influenced the way of thinking and research in the humanities [13]. More precisely, the early digital humanities were mainly concentrated in the fields of philosophy, language, and literature [14], and then gradually expanded to history, geography, religion, art and other fields [15]. For the vast majority of humanities research, text data is the most important type of data [16], so text analysis has become the basic method of digital humanities [17]. In addition, other DH tools, technologies, methodologies, theories and practices are widely studied by researchers in the whole word.

Many review studies on digital humanities are helpful for us to accurately understand the progress and current situation in this field. Qing, Wang [18] based on 803 records collected from the web of science database, which were related to DH research from 1968 to 2017, and using VOSviewer and CiteSpace visual analysis tools, summarized 4 main research topics in DH field: (1)library and information services for digital humanities project; (2)digital history; (3)digital literary; and (4)digital cultural heritage. Similarly, Fang, Su and Yin Zhang [19] also adopted the bibliometrics method, combined with social network analysis and visualization tools for the longitudinal examination, and finally identified 4 major research topics in the DH field: (1) collections and contents; (2)technologies, techniques, theories and methods; (3)collaboration, interdisciplinarity and support; and (4) DH evolution. Gupta, Nidhi and Chakravarty, Rupak [20] tried to find the current state of the academic literature regarding DH by using scientometric analysis, and they identified the highly contributing researchers, countries, organizations, and main subject areas of DH. Chansanam, Wirapong et al. [21] also conducted a scientometric analysis, in addition to the basic bibliometric information of publishers, authors, citations, institutions, and countries, they focusing on the structure of knowledge of DH, including famous themes, co-citations, and bibliographic networks. Moreover, Muh-Chyun, Tang et al. [22] did a bibiometric analysis of the literature published in digital humanities, finding that limited by language and geographic boundaries the network of co-authors is still quite fragmented.

These reviews show the current research status of digital humanities from different perspectives around the world. However, according to our observation, few scholars have deeply focused on the research of digital humanities in a certain region. With a long history and rich cultural heritage, China has been the cradle of DH academic achievements in recent years and therefore should not be underestimated in the DH field. Xiaoguang, Wang et al. [23] reviewed the history and present situation of Chinese DH scholarship by means of historical analysis and bibliometrics. Taiwan has made many achievements in digital collection and digital humanities, Summarizing the theoretical research and practice progress of DH in Taiwan can show the achievements of digital humanities in a specific region and provide reference for other regions to carry out DH research. This study will provide researchers and scholars with the latest trends and developments in the field of DH research in Taiwan.

3. Methodology

3.1. Data collection

As discussed above, after the accumulation of 7 conference, Taiwan digital humanities research has entered a new stage since the 8th DADH in 2017, which begins to seek the balance between digital technology and humanistic meaning, as described in the website of the DADH2017 Conference:

> "When the research tools and topics of interest become common, how far can digital humanists and social scientists collaborate to gain deeper insights into common research problems and cope with shared challenges?"[DADH2017. https://www.aiecon. org/conference/DADH2017/index.htm]

Therefore, we take 2017 as the starting point and choose 2017–2021 as the research time range. We obtained the conference proceedings from the official website of the conference, a total of 217 conference papers, which constitute the main content of our research. We obtained the conference proceedings from the official website of the conference, a total of 217 conference papers, which constitute the main content of our research. *The Journal of Digital Archives and Digital Humanities* is an additional product of the DADH since 2018, some papers of the conference have been published in the journal at the same time. Therefore, we use the journal papers as supplementary materials. We searched by journal name *The Journal of Digital Archives and Digital Humanities* from the *Airitilibrary* online database (http://www.airitilibrary.cn) and obtained 35 journal literatures, and a total of 252 papers were obtained. The number of conference and journal papers each year is shown in Table 1.

Referring to the idea of three-level coding of grounded theory, the article is regrouped by means of manual interpretation. According to the content information conveyed by the title, abstract and text of the article, and the category of the article in the conference or journal, the 252 articles are classified according to the content, mainly divided into three categories: 1) viewpoint and theory, 2) tools and techniques, and 3) practice in various fields. The secondary classification and description of the three categories is shown in Table 2, and the article content will be mined and analyzed according to this classification.

The division principle is as follows: for example, "Digital Ricoeur: a digital research portal" in the conference in 2017 [24], as described in the passage "an ongoing project of providing digital research tools", accordingly this article is deemed to belong to tools and techniques. Another example is the category of "Digital Humanities infrastructure and tools" set up in the 10th International Conference of Digital Archives and Digital Humanities [25]. According to this, it is judged that the four articles under this category are all about tools. For another example, the five articles in the 12th International Conference of Digital Archives and Digital Humanities belong to the category of "social network analysis" [26]. Social network analysis is one of the commonly used research methods in digital humanities, and according to the content conveyed by the article title, the five articles are deemed to belong to the practice category. According to the above principle, all 252 articles were manually interpreted and classified.

3.2. Paper statistics

There are 37 articles in the category of Viewpoint and Theory, 86 articles in the category of Tools and Techniques, and 129 articles in the category of Practice. The specific number distribution in each year and the specific breakdown under the three categories are shown in Figs. 1 and 2 respectively. Since the number of papers included in each conference is not fixed, the trend in the figure cannot represent the changing trend of research focus. However, from the proportion of the number of papers, we can see that the number of papers in the category of Practice accounts for the most, the number of papers in the category of Tools and Techniques is second, and the number of theoretical research papers is the least. In addition, from the perspective of secondary classification, the number of articles in the text tools category is the largest among all categories, corresponding to the number of practice papers in the field of literature also occupying the highest proportion. This phenomenon shows that the research of digital humanities in Taiwan focuses on the development and technical discussion of text tools, and makes full use of text tools in the practice of digital humanities of literary materials. This is because the research materials of digital humanities are usually digital text materials, so a large number of text-related research results are produced.

3.3. Word frequency statistics

Further, take word frequency statistical analysis on the title of the article and the keywords of the article itself. First, all the article titles and keywords are unified into simplified Chinese, and proper nouns such as AR, GIS, OCR, and AI are reserved. Then, stop-words such as "to", "of", "based on", "and" and words without special reference such as "research", "application", "method", "take … as an example", "thinking", "development" and "prospect" are removed from the title through manual interpretation. At the same time,

Table 1
The number of papers in each year.

	2017	2018	2019	2020	2021	All
Conference papers	54	30	28	34	71	217
Journal papers	0	11	8	8	8	35
All	54	41	36	42	79	252

Table 2
The description of the three categories.

| category	secondary classification	description
Viewpoint and Theory	Basic theory	Research on the pure theory of digital humanities.
	the Perspective of Libraries Museums and Archives	The discussion of digital humanities from the perspective of libraries, museums and archives.
	Expansion of Research Perspective	In what fields can digital humanities be applied in the future.
	Critical Thinking	Critical thinking and discussion triggered by digital humanities.
Tools and Techniques	Text Tools	For text analysis of various types of word segmentation, marking, statistics and other tools development and algorithm application.
	Network Analysis Tools	Introduction to the development and use of research tools based on social network analysis.
	GIS/Visualization Tools	GIS system, image, video, AR, VR and other visualization technologies
	Integrated System Platform	Such as DocuSky, CBDB and other comprehensive system, platform, database, corpus.
Practice	Literature	The practice of digital humanities with various literary works as the research object.
	History	The practice of digital humanities from the perspective of history.
	Linguistics	Practice in linguistics with an approach of digital humanities.
	Religion	The practice of digital humanities involving religions such as Buddhism.
	Social	The practice of digital humanities from the perspective of social science.
	AI	The practice of combining artificial intelligence and digital humanities.
	Myth	The practice of digital humanities based on fairy tales.
	Art	The practice of digital humanities with art works such as painting and music as research objects.
	Politics	The practice of digital humanities from the perspective of politics.
	Others	The practice of digital humanities in other fields, such as film, song lyrics, journalism and law, are listed in the "Other" category because there is only one relevant article.
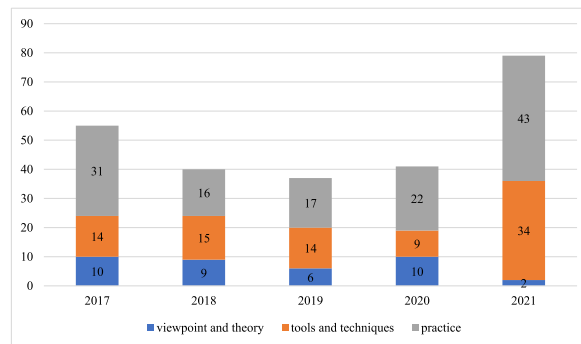
Fig. 1. Statistics on the number of three categories of articles from 2017 to 2021.
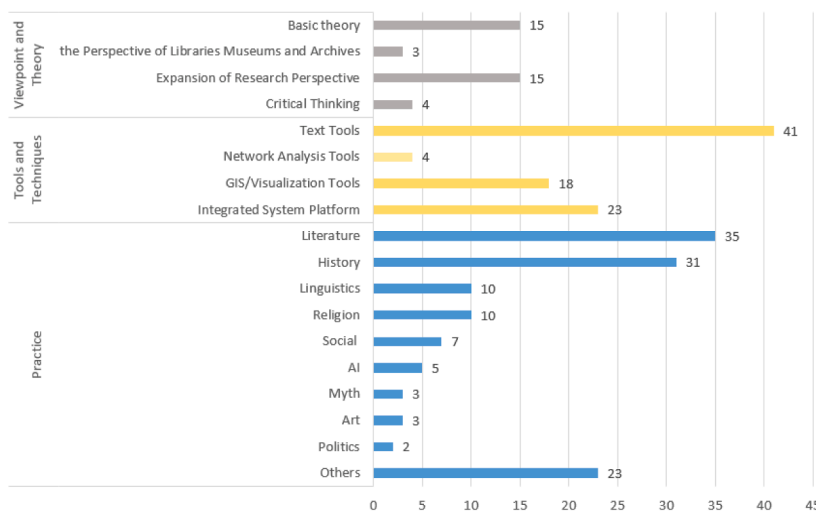

Fig. 2. Numbers of articles under three categories.

word segmentation is performed on the title to form a keyword list of 1464 words. Use Python to perform word frequency statistics on the keyword list.

Since this research is conducted under the theme of digital humanities, the words such as "digital humanities" and "digital humanities research" in the keyword list have no special significance to the research, so they are excluded. In addition, similar words such as "text", "text exploration", "text analysis", "social network", "network analysis" and "social network analysis" are manually merged. As shown in Table 3, 19 keywords with word frequency ≥5 were finally obtained.

In the statistical results of word frequency, "text" has the highest frequency, which is consistent with the phenomenon that there are many text tools and practical articles in the field of literature in the above-mentioned classification of papers. High-frequency words such as "tools", "DocuSky", "corpus", "database" and "visualization" have obviously technical orientations, showing that digital humanities research in Taiwan focuses on the technical area. High-frequency words such as "history", "personage" and "biography" have significant theme characteristics, which indicates that digital humanities research in Taiwan focuses on terms of content. Among them, the word "Taiwan" also occupies a high position, indicating that more attention is paid to local culture in practice. The high-frequency word "teaching" means that Taiwan has integrated digital humanities into teaching and opened relevant courses.

4. Analysis and discussion

4.1. Viewpoint and theory

Theoretical research is the cornerstone of supporting the development of digital humanities. There are 37 articles on digital humanities theory in the research sample. Although it accounts for less than 15% of the total amount, the analysis of its content can also roughly describe the research focus on the theory of digital humanities in Taiwan. From the distribution of the number of papers, many articles are about the research and discussion of pure theory and the expansion of the possible research field of digital humanities. In addition, a few articles put forward the development prospect of libraries, museums and archives from the perspective of digital humanities, and some articles critically consider the problems that may arise under the development of digital humanities.

4.1.1. Basic theory

Digital humanities refer to the use of modern information technology to solve research problems in the field of humanities. In other words, before the emergence of digital humanities, many research questions were difficult to observe and imagine, and the emergence of digital humanities provided new solutions to these problems based on computational thinking. Computational thinking is the basis for digital humanities to use digital tools to innovate the theory, concept and method of humanities research [27]. In other words, computational thinking provides a logical context and a framework for abstract thinking for research, while digital humanities is a specific research topic and application scenario, which is the practice and verification of the research framework under computational thinking [27]. A large number of digital research materials are the premise of digital humanities research. Digital humanities research in Taiwan is also based on digital databases. The construction of the database should be carried out from the perspectives of data sharing and member communication [28]. Not only digital data sharing but also include the sharing of various tools, research methods and technologies [28]. Member communication includes communication among university institutions, research teams, libraries, publishing houses, commercial database companies, foundations, etc [28].

In the face of massive digital resources, Text analysis is a part or even an important part of digital humanities [29]. The most significant contribution of digital technology to humanities research is retrieval function and word frequency analysis. In addition, the main contents of text analysis under digital humanities include relevant texts and research context, Chinese poetry analysis and semantic changes, design and collaborative application of analysis tools [29]. Under the open-data environment, different researchers can observe the same research content from different perspectives, to enhance the dialectics of digital humanities research [29].

4.1.2. The perspective of libraries museums and archives

Libraries, museums and archives are the main units of digital collections that have the resource advantages of digital humanities, and play an important role in digital humanities. In this context, libraries are not only a provider of various paper and digital resources, but also should gradually build up databases and tools, become the creator of the databases and the developer of research tools, and participate in various digital humanities research projects to a certain extent, and also be a digital humanities researcher [30]. Museum

Table 3
Word frequency statistics.

word	frequency	word	frequency
text analysis	34	visualization	12
history	20	personage	11
tools	20	museum	8
social network analysis	19	picture	8
DocuSky	18	biography	8
Taiwan	16	digitization	6
teaching	15	space	5
corpus	15	AI	5
database	14		

collections contain important cultural values and play a role in spreading culture to the public. Digitization of collections provides important opportunities for cultural communication. In the process of gradually promoting the digitization of museum collections, policies, rules and regulations should also be considered, especially how to stimulate the enthusiasm of the public to enhance the participation and contribution of the public under the premise that the existing work is led by the government [31]. The rapid development of digital technology has promoted the archives industry to gradually shift from the physical management of archival materials to content management. The current research on archival content management mainly focuses on the specific technology or system construction and lacks the overall analysis at the global level [32]. Therefore, a complete archival content management model should be built.

4.1.3. Expansion of research perspective

Digital technologies have expanded a new perspective for research in the field of humanities. For example, Nurjanah Jane and Inoue Hiroki [33] proposed that under the digital humanities perspective, the tsunami in Aceh, Indonesia can be reproduced through digital archives and other channels; Deng Shuyang [34] took the "engraving printing" technology in intangible cultural heritage as the research object and proposed a digital protection scheme; Ming Yeung Cheung and Yi-Fu Chen [35] tried to mark religious documents through TEI, with a view to realizing the progress from "religious research" to "digital religious research". These are places that have received little attention in the past research and practice, and even have never been practiced before. It can be said that the emergence of digital humanities makes the research burst into new vitality.

4.1.4. Critical thinking

Different from the ideas and methods of traditional humanities, digital humanities lays more emphasis on the application of digital technologies. These digital technologies certainly provide many conveniences for the research of humanities, but the discussion on the two-sided nature of technologies has never stopped. Because digital humanities focus on digital data mining, they lost the traditional humanities analysis of rich subjective experience [36]. Moreover, although digital technologies are included in the research of humanities, it does not mean that humanities scholars can rely excessively on "Computational Thinking" and should continue to maintain the textual research on data. This literacy can be called "Digital Textual Criticism" [37]. At the same time, digital humanism is accompanied by the generation of massive digital data, this phenomenon will inevitably cause a new form of digital privacy infringement risk to citizens, and it is urgent to formulate a standardized system [38].

4.2. Tools and techniques

There are 86 papers on tools and techniques. According to the distribution of papers, 41 papers are text tools, accounting for the highest proportion. In addition, the research on tools and techniques of digital humanities in Taiwan mainly focuses on the establishment of a comprehensive platform system, GIS technology, image recognition technology, visualization, artificial intelligence, and social network analysis tools. There are also discussions on metadata, linked data, open data, and knowledge ontology. However, there are few papers on these topics and no research hotspot. This section mainly discusses text tools on digital humanities, network analysis technology, GIS/visualization, and comprehensive platform system in Taiwan.

4.2.1. Text tools

At present, more and more original documents have been digitized, but most of them are unstructured data. Since various computer technologies are only designed for structured data, so natural language processing technology needs to be adopted to realize the structured processing of research data [39]. Then, research data is further processed by using the techniques of word segmentation, marking and content reading. Liu Huiwen [40] discussed the lexicon setting of commonly used CKIP and Jieba word breakers from the perspective of pragmatics, and analyzed the possible result deviation. Alvin Cheng-Hsien Chen's team developed an information marking system for biographical texts [41]; Wai-Him Pang [42] introduced a computer-aided full-text marking and information semi-automatic extraction system developed for Chinese Local Chronicles. By extracting the key features such as name, time and official position in the data, it helps scholars to efficiently obtain the characteristic information of "who, when and any officeholder". Keywords play an important role in text mining, which not only help researchers to quickly insight the content of text but also serves as an important basis for understanding the style and grasping the spatial and temporal background of the text. Therefore, keyword acquisition is the first and extremely important step of text mining. Ching-Syang Jack Yue [43] introduced the common method of keywords detection: TF-IDF (Term Frequency–Inverse Document Frequency) at the conference in 2018, and proposed that the keywords accuracy of TF-IDF is affected by factors such as the text and the number of words in papers, and statistical thinking can be added to improve the accuracy. Subsequently, at the conference in 2019, Ching-Syang Jack Yue and Xu Cheng-en [44] proposed an unsupervised learning model for keywords acquisition with t-SHE as the core technology, and identified potential keywords in each word frequency through text structure, data grouping, vocabulary detection and other processes.

4.2.2. Network analysis tools

People are the key elements of history. The analysis of social networks around typical people of different social strata and historical periods promotes to discover the complex relationship of characters hidden in historical materials. It is a significant content of historical research. In addition, biographies, official records, genealogies, novels and other materials all have typical character features. When conducting digital humanities research on such texts, network analysis among characters is also a major topic. Therefore, in the establishment of social networks, Xie Shunhong [45] proposed that manpower and computers should be combined, and the concept of

"common characteristic value" should be used to give specific concepts or characteristics, such as community and residence, to guide the computer to automatically merge the social networks of multiple people, and then establish the social relationship network in the way of automatic hierarchical grouping. In this way, the consumption of labor costs can be avoided, and it can also ensure that the automatically established network has a meaningful association.

For the presentation of the results of network analysis, Chen Zhiming and Zhang Zhong [46] constructed the "historical data character relationship map tool". Firstly, the tool preprocesses text word segmentation, extracts the name information and establishes the character social relationship matrix, then establishes a visual social network analysis map, and provides researchers with an interactive interface for editing the characters' relationship. At the same time, Chen Zhiming and Zhang Zhong [46] proposed that the social network analysis tools developed around the historical characters' relationship map in the future, should also provide social network measurement and analysis of kinetic energy including centrality, structural holes and the most influential nodes, and can be arbitrarily matched with different collections to analyze the characters relationship, to more effectively assist researchers in mining the hidden relationship context in the text.

4.2.3. GIS/visualization tools

The combination of historical content and geographical concepts is a common phenomenon in the study of historical geography. GIS has also become a significant part of digital humanities. GIS-based digital humanities research is according on the information of time and spatial of the research data, extracts information such as age, dynasty, region and area etc., and finally, the spatial structure of historical information is reproduced in the form of an electronic map, so that researchers can more intuitively feel the historical dynamic changes in spatial coordinates. Lin Nongyao [47] carried out information processing on historical place names in Taiwan and spatial position, combined with the historical and cultural map resources of Taiwan Academia Sinica, and constructed an API interface of historical place names in Taiwan that can be used by other platforms or programs. Finally, it formed a place name query platform for the whole Taiwan region, so that researchers could obtain more complete place name query functions.

With the growing development of GIS, AR, VR and other visualization technologies, the discussion on various visualization technologies in the research practice of digital humanities is also deepening. Shih-Chueh Kao and Vince Briffa [48] combined picture data with AR technology to visually reproduce prehistoric buildings; Brian Kokensparger [49] proposed a visual research scheme for the path of characters in novels, and visually analyzed the characters in 38 plays of Shakespeare to build the footprints of the characters in the plays; Tai-Jui Wang [50] tried to use 3D capture technology to explore the dynamic visualization of Chinese opera.

The above social network analysis often needs to present the relationship network map of the research results in a visual manner. More scholars have proposed a new perspective of social network analysis based on image recognition technology. Biographical materials also contain a large number of old photos such as personal photos, family photos and group photos. Discoveries can often be made by identifying the characters in these group photos and combining specific factors such as photo location and filming time. So, Zhang Sufen [51] used face recognition to analyze the social network of the group portrait in the old photos based on the people data in the "Taiwan biographical database (TBDB)".

4.2.4. Integrated system platform

According to the statistical results, 23 papers are the research and introduction of comprehensive digital humanities platforms, systems and databases, such as CBETA digital research platform for buddhist research [52]; the graphic database [53] constructed by taking "Yupian" in Song Dynasty as the research object. In particular, 8 of them are about the DocuSky platform [54–61], occupying a considerable weight. Hsieh-Chang Tu [54], one of the platform developers, pointed out that the text database established by institutions may not meet the needs of researchers, and there should be a mechanism for researchers to build their own personal database. Under this motivation, the Digital Humanities Research Center and the Digital Collection and Automatic Inference Laboratory of the Department of Information Engineering of National Taiwan University jointly developed the DocuSky platform. The platform also provides text tools for database management, retrieval and browsing, analysis, geographic information presentation, and so on. With these digital tools, users can conduct more in-depth research and analysis the text they are interested in. At present, DocuSky has become a comprehensive digital humanities network platform integrating data processing and various analysis tools, and is an important tool for digital humanities research in Taiwan.

4.3. Practice in various fields

Digital humanities practice based on actual cases is the main form of digital humanities research in Taiwan. There are 129 papers of this kind in the research sample, more than half of all papers, and it is the most numerous category. Literature is the main field of digital humanities practice in Taiwan, which is also matched with the result that text tools occupy a dominant position in the previous discussion, and the field of history, language and society is also popular area of digital humanities practice in Taiwan.

It is observed that the digital humanities research in Taiwan attaches great importance to the religious field mainly based on Buddhism. There are not only tools and technologies specially developed for Buddhist scriptures [62,63], but also lots of practice in many religious fields. There are 10 papers involving religion, especially Buddhist scriptures such as the Tibetan Sutra [64–68], and research on Buddhist grottoes [69] and Christianity [70]. In addition, with the continuous development of digital humanities in the religious field, the 9th. International Conference of Digital Archives and Digital Humanities in 2018 set up the section of "Digital humanities research on Buddhist literature". In 2020, the theme paper report of "Digital humanities in the Religious Context" was also launched.

4.3.1. Literature field

There are 35 papers in the field of literature. The practice in this field mainly uses Tang poetry, the Book of Songs, the Analects of Confucius, novels, biographies and other literatures as research materials, and applies a large number of methods and technologies of word frequency statistics, content analysis and social network analysis.

The research in the literary field of digital humanities can not be separated from word frequency statistics. The quantitative method can be used to analyze the language characteristics of the text and the writing habits of the author. Shih-Wen Chyu [71] took the thirty aristocratic texts of "Historian Records" as the object, made word frequency statistics on function words in the text, so as to analyze the characteristics of different authors, and proves the author's attribution of " Records of the Historian" aristocratic chapters. Similarly, Ching-Syang Jack Yue and Yu-Ting Ye [72] analyzed the reporting styles of Taiwan's four major newspapers by studying the application of vocabulary, sentence length, punctuation, and function words; Liao Xueying [73] made quantitative statistics on the words used in the "Book of Songs", and studied the microstructures such as rhyme, syntax and repetition in the poems, to capture the rhythm form and language characteristics of the "Book of Songs".

In addition, text content analysis is a major focus in the field of literature, mainly from the perspective of keyword analysis, co-occurrence words and emotional words analysis, to achieve a deep-seated study of a certain theme of the text. For example, Zheng Wenhui [74] conducted keyword analysis on the "Tang Poems" and found that the word "where" was used most frequently, especially in the middle and late Tang Dynasty, "where" accounted for more than half of the poems. Further through the analysis of co-occurrence words and emotional words, combined with the characteristics of the times and other factors, it is found that the essence is the change in the emotion of poets caused by the social reform in the middle Tang Dynasty, which makes the poets feel confused, lost, hesitant and other feelings without belonging in reality, then cast them into the poems to become the question of "where".

4.3.2. History field

Digital humanities technologies provide a new perspective for the study of historiography. The help of digital technologies, not only makes the digital restoration and collection of a large number of historical relics possible but also enables large-scale historical-geographical research from a macro perspective. The materials of historical research are very broad, involving population, language, folklore, culture, politics, economy, climate and other aspects, digital humanities practice can be carried out from any angle to generate new knowledge discovery. For example, Lin Ma and Monica Li [75] analyzed the promotion path of officials in the Song Dynasty; Xu Yahui [76] revealed the burial culture and historical changes in the Song Dynasty by analyzing the regional distribution and variety of the tombs in the Liao, Song and Jin Dynasties.

4.3.3. Other fields

Language, society, mythology, art, politics, and other fields are the areas where digital humanities research in Taiwan is concentrated. For example, Ye Qiuxing [77] analyzed the vocabulary use characteristics of the Hakka dialect from the perspective of linguistics according to the Hakka language database; Ming Ren [78] studied the phenomenon of "Beijing Drifters" from the perspective of sociology and explored the collective memories in the digital age.

For another example, Wang Ping [79] selected more than 120 paintings as experimental samples and another 240 paintings from the South Song dynasty as supporting materials, using image capture technology on paintings with two easily recognizable motifs 'fish-boat' and 'fisherman', and studied the generation and change of the theme of " Fisherman-Hermit " in the paintings. It is found that the development of the cultural symbol "Fisherman-Hermit" prevails in poetry and painting after the Southern Song Dynasty, which is related to the rise of Neo-Confucianism in the Song Dynasty. In the middle and late Southern Song Dynasty, especially in the Yuan Dynasty, due to the reduction of the political status of poets, "Fisherman-Hermit " as the subject or theme of poetry and painting was more common. This research fully shows the value of digital humanities for discovering new tacit knowledge, especially for sorting out and revealing the potential context in history and culture. Through the methods of digital humanities, unexpected discoveries can often be made.

There are some interesting practices and attempts in the research of digital humanities in Taiwan beyond the scope of conventional academic discussions. For example, Wei-Min Fan and Muh-Chyun Tang [80] took the nominees of the Best Chinese Male and Female Singer Awards of the 18th to 29th Taiwan Golden Melody Awards as the research samples, and explored whether the nominees were further involved in song creation in addition to acting as singers. Finally, they found writing lyrics was the main thing of singers that participated in the creation of songs. Not only that, the "Classic of Mountains and Rivers", rap lyrics and other contents are also the materials for digital humanities research in Taiwan. These practices and attempts can be regarded as an effective means to broaden minds and open the horizon of researchers.

5. Conclusion

After more than ten years of continuous development, the digital humanities in Taiwan have formed a certain scale. Many research results and practical experiences have reference significance for digital humanities research in mainland China. Based on the research of the International Conference of Digital Archives and Digital Humanities in Taiwan, 252 papers are analyzed and studied by means of econometric statistics, text analysis, theme analysis and other research methods. It is believed that the digital humanities research in Taiwan presents the following characteristics.

(1) The digital collection work led by the government has made a large number of materials digitized, which has laid a solid resource foundation for the follow-up digital humanities research. Therefore, the digital humanities in Taiwan started from

practice. The number distribution of samples also confirms this phenomenon, that is, in sharp contrast to the extensive digital humanities practice, the research on digital humanities theory is relatively weak.

(2) The vast majority of digital humanities research databases are text-based literature. For text, text mining and text analysis are the key research methods, and it is particularly important to develop various text research tools around this. This objective fact has led to the development of a large number of text analysis tools such as punctuation, tagging, content acquisition, text reading, keyword detection, and so on.

(3) Taiwan has continuously and deeply carried out digital humanities practice in various fields. In addition to the literature and history fields, which were widely involved by traditional academic circles, Taiwan focuses more on the religious field and local culture, and has expanded many interesting practices and attempts, such as the research on Golden Melody Awards and rap lyrics, which provides important ideas for other scholars to start a digital humanities research in the future.

In general, digital humanities have the characteristics of interdisciplinary, while humanities scholars often do not have a high-tech level. Therefore, developing a common tool or platform for researchers in the general humanities field and reducing the threshold of digital humanities research are bound to become the key issues to be solved in the future. In addition, most of the research is directly engaged in the concrete practice of digital humanities at present, while pure theory research is scarce. The development of digital humanities without theoretical support will lack soul and internal driving force. Therefore, future digital humanities research should focus more on the promotion and deepening of basic theories. The research in the field of humanities carries the inheritance of culture. With the help of digital humanities, we can not only explore new knowledge in the field of humanities but also enhance the availability of knowledge, vividly display cultural knowledge and promote cultural promotion and dissemination. Therefore, it can be predicted that expanding the promotion and dissemination of research results through various visual experience technologies and undertaking the task of cultural inheritance and promotion will become an important cultural mission of digital humanities research in the future.

Author contribution statement

All authors listed have significantly contributed to the development and the writing of this article.

Data availability statement

Data included in article/supp. Material/referenced in article.

Additional information

No additional information is available for this paper.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

References

[1] Hsieh-Chang Tu, Jieh Hsiang, Taiwan history digital library (THDL), in: Research Center for Digital Humanities of National Taiwan University, 2009, https://doi.org/10.6681/NTURCDH.DB_THDL/Text.

[2] National Palace Museum, Research center for digital humanities of national taiwan university, in: National Palace Museum Digital Library of Qing Archives. National Palace Museum, 2022.

[3] Research Center for Digital Humanities of National Taiwan University, Context discovery system for lidaibao'an, in: Research Center for Digital Humanities of National Taiwan University, National Taiwan University Library, 2020. https://airiti.com/LandingPage/NTURCDH/10.6681/NTURCDH.lidaibaoan/Text.

[4] DADH. http://www.dadh-record.digital.ntu.edu.tw/Scope.php?LangType=tw&His=D09k2, 2009.

[5] DADH. http://www.dadh-record.digital.ntu.edu.tw/Scope.php?LangType=tw&His=Y10k2, 2010.

[6] DADH. http://www.dadh-record.digital.ntu.edu.tw/Scope.php?LangType=tw&His=MMXXI, 2011.

[7] DADH. http://www.dadh-record.digital.ntu.edu.tw/introduction.php?LangType=tw&His=MKOOO, 2012.

[8] DADH. http://www.dadh-record.digital.ntu.edu.tw/Agenda.php?LangType=tw&His=EKALM, 2014.

[9] DADH. http://dadh-2015.digital.ntu.edu.tw/ch/conference-program, 2015.

[10] DADH. http://www.dadh-record.digital.ntu.edu.tw/Scope.php?LangType=tw&His=BHUGF, 2016.

[11] DADH. https://www.aiecon.org/conference/DADH2017/index_ch.htm, 2017.

[12] Susan Schreibman, Ray Siemens, John Unsworth (Eds.), A Companion to Digital Humanities, Blackwell, Oxford, 2004. http://www.digitalhumanities.org/companion/.

[13] Willard McCarty, Humanities Computing, Marcel Dekker Press, New York, 2003.

[14] Joseph Raben, Humanities computing 25 Years later, Comput. Humanit. 26 (6) (1991) 341–350.

[15] Willard McCarty, Harold Short, Mapping the Field, European Association for Digital Humanities, 2002. http://eadh.org/node/188.

[16] S. Mahony, Cultural diversity and the digital humanities, Fudan J. Humanit. Soc. Sci. 11 (3) (2018) 371–388.

[17] G. Rockwell, S. Sinclair, Hermeneutica: Computer-Assisted Interpretation in the Humanities, MIT Press, Cambridge, MA, 2016.

[18] Q. Wang, Distribution features and intellectual structures of digital humanities: a bibliometric analysis, J. Doc. (2018), https://doi.org/10.1108/JD-05-2017-0076.

[19] F. Su, Y. Zhang, Research output, intellectual structures and contributors of digital humanities research: a longitudinal analysis 2005–2020, J. Doc. (2022), https://doi.org/10.1108/JD-11-2020-0199.

[20] N. Gupta, R. Chakravarty, Science Mapping Analysis of Digital Humanities Research: A Scientometric Study, Library Philosophy and Practice, 2021.

[21] W. Chansanam, A.R. Ahmad, C. Li, Contemporary and future research of digital humanities: a scientometric analysis, Bullet. Elect. Engin. Inform. 11 (2) (2022) 1143–1156, https://doi.org/10.11591/eei.v11i2.3596.

[22] M.-C. Tang, Y.J. Cheng, K.H. Chen, A longitudinal study of intellectual cohesion in digital humanities using bibliometric analyses, Scientometrics (2017), https://doi.org/10.1007/s11192-017-2496-6.

[23] X. Wang, X. Tan, H. Li, The evolution of digital humanities in China, Libr. Trends (2020), https://doi.org/10.1353/lib.2020.0029.

[24] Ming Yeung Cheung, Digital Ricoeur: a digital research portal, in: 8th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2017, pp. 23–26.

[25] Program for the 10th international conference of digital archives and digital humanities[EB/OL]. [2022.05.23]https://dadh2019.conf.tw/site/page.aspx?pid=299&sid=1308&lang=CHT#paper3.

[26] Program for the 12th international conference of digital archives and digital humanities[EB/OL]. [2022.05.23]http://dadh2021.ncue.edu.tw/pages/program.html.

[27] Jyi-Shane Liu, Computational thinking and implications for digital humanities research, J. Digit. Arch. Digit. Humanit. 1 (2018) 51–77.

[28] Hongsu Wang, Lik Hang Tsui, K. Peter, Bol. A cyberinfrastructure for historical China studies, J. Digit. Arch. Digit. Humanit. 6 (2020) 1–35.

[29] Chao-Lin Liu, Toward linking open data, synchronizing digital tools, and mapping relevant topics: a survey of the series on digital humanities, J. Digit. Arch. Digit. Humanit. 1 (2018) 11–50.

[30] Fengen Tu, Lixuan Yang, Rethinking on open tools, digital humanities and library, in: 9th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2018, pp. 38–54.

[31] Hui-Tuan Chang, Public practice of anthropological collections (archives) and its digitalization in museums of taiwan, in: 8th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2017, pp. 89–91.

[32] Xuemin Guo, Trend analysis of digital archives management in big data era: based on 5W1H method of content management, in: 8th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2017, p. 118.

[33] Nurjanah Jane, Hiroki Inoue, Hidenori Watanave, Aceh tsunami recurrence from historical records and local interviews, in: 8th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2017, pp. 86–87.

[34] Shuyang Deng, Huaidong Ge, Tie Shu, Study on digital protection of intangible cultural heritage "engraving printing", in: 9th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2018, pp. 518–533.

[35] Ming Yeung Cheung, Yi-Fu Chen, From religious studies to digital religious studies: experiences and reflections for TEI tagging on journal articles of religious studies, in: 10th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2019, pp. 220–244.

[36] C. Paolo, A Critical Approach to Digital Humanities: Examining the Subjective Sphere, vol. 2, 2018, pp. 1–13.

[37] Fu-shih Lin, Digital textual criticism: the new literacy for humanities scholars, J. Digit. Arch. Digit. Humanit. 5 (2020) 1–35.

[38] Hsiao Yu Chung, Big data and civil society: from the perspective of digital privacy, in: 8th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2017, pp. 95–97.

[39] Dongbo Wang, Wenhao Ye, Jiwen Liang, The application of natural language processing technology in digital humanities, in: 8th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2017, p. 65.

[40] Hui-Wen Liu, The meta- and theoretical meaning of automatic Chinese segmentation: a pragmatism perspective, in: 8th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2017, pp. 66–68.

[41] Alvin Cheng-Hsien Chen, Táňa Dluhošová, Taiwan biographical ontology (tbio): a working annotation scheme for biographical information extraction and its application to sociology of literature, in: 8th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2017, p. 117.

[42] Wai-Him Pang, Hui Cheng, Shih-Pei Chen, Toward linking open data, synchronizing digital tools, and mapping relevant topics: a survey of the series on digital humanities, J. Digit. Arch. Digit. Humanit. 1 (2018) 79–125.

[43] Ching-Syang Jack Yue, Chengen Xu, Yingyi Liang, Comparison and application of keywords extraction methods, in: 9th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2018, pp. 428–462.

[44] Ching-Syang Jack Yue, Chengen Xu, A statistical approach of keywords extraction for Chinese documents, in: 10th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2019, pp. 2–20.

[45] Shunhong Xie, Haoren Ke, Sufen Zhang, The application of automatic clustering in the construction of relationship between historical people, in: 9th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2018, pp. 125–135.

[46] Zhiming Chen, Zhong Zhang, Development and application of an ancient books digital humanities research platform with characters' relationship map tool, in: 9th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2018, pp. 173–203.

[47] Nongrao Lin, Haoyang He, Jun lin Guo, Integration and development of historical gazetteer system in Greater China, in: 9th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2018, pp. 204–220.

[48] Shih-Chueh Kao, Vince Briffa, Revelation of prehistoric architecture through augmented reality: representation of the sanctuary at tas-silġ, in: 8th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2017, pp. 73–75.

[49] Brian Kokensparger, Walking into a brave new world: visualizing character paths through shakespeare's plays, in: 8th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2017, p. 72.

[50] Tai-Jui Wang, Research on the teaching and learning of using 3D motion capture analysis on Chinese opera performance, in: 8th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2017, pp. 76–77.

[51] Sufen Zhang, Shunhong Xie, Haoren Ke, Group photos in TBDB: another way to explore the social network, in: 10th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2019, pp. 214–219.

[52] Jen-Jou Hung, CBETA research platform: a digital tool for studying Chinese buddhist texts in the new era, J. Digit. Arch. Digit. Humanit. (1) (2018) 149–174.

[53] Guanwei Liu, The construction of the graphic database of the song edition of "yupian", in: 10th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2019, pp. 118–126.

[54] Hsieh-Chang Tu, DocuSky and its applications to the visualization of text-term relationship graph, in: 9th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2018, pp. 2–20.

[55] Hsieh-Chang Tu, DocuSky: a platform for constructing and analyzing personal text databases, Journal of Digital Archives and Digital Humanities 2 (2018) 71–90.

[56] Chi-jui Hu, Hsieh-Chang Tu, Cheng Chen, Enriching the digital texts through the DocuSky metadata attachment tool, J. Digit. Arch. Digit. Humanit. 4 (2019) 71–107.

[57] Kuan-Lin Chen. The porting and significance of veritable records of the joseon dynasty on the DocuSky collaboration platform[EB/OL]. [2021.11.12]. https://dadh2020.ascdc.tw/#dadh-program.

[58] I-Mei Hung, Te-Chi Tsao, Design and application of metadata in digital humanities basing on DocuSky: a case study of buddhist inscriptions, J. Digit. Arch. Digit. Humanit. 5 (2020) 117–153.

[59] Hsieh-Chang Tu, Binary correlation visualization tool on DocuSky[EB/OL] [2022.05.24], http://dadh2021.ncue.edu.tw/pages/program.html.

[60] Huang Jiahong. The establishment of docusky online public database on unified records of the ming dynasty[EB/OL]. [2022.05.24]. http://dadh2021.ncue.edu.tw/pages/program.html.

[61] Cao Deqi. Text oriented DocuSky platform teaching[EB/OL]. [2022.05.24]. http://dadh2021.ncue.edu.tw/pages/program.html.

[62] Xianchao Shi, A method and implementation of automatic punctuation, J. Digit. Arch. Digit. Humanit. 3 (2019) 1–19.

[63] Jen-Jou Hung, L.I. Chin-Hsien, A requirement analysis and the implementation and evaluation of an online parallel corpus editing tool based on research needs in buddhist textual studies, J. Digit. Arch. Digit. Humanit. 8 (2021) 133–185.

[64] Guanwei Liu, Naiqi Li, Constructing a full-text database of "pronunciation and meaning of all the scriptures (yiqiejing yinyi)":A collection of ten ancient manuscripts in Japan[C], in: 8th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2017, p. 79.

[65] Marcus Bingenheimer. Chinese buddhism as historical social network(s)[EB/OL]. [2022.05.24]. http://dadh2018.dila.edu.tw/agenda/.

[66] Jen-Jou Hung, Tian-Hua SHI, Huei-Chan SHI. Extraction and analysis of the citation of Chinese buddhist electronic texts in buddhist studies publications. [EB/OL].[2022.05.24]. https://dadh2020.ascdc.tw/#dadh-program.

[67] Shu-Ling Huang, Yu-Chun Wang. The study of ontology in digital humanities– Taking the construction and application of Buddhist ontology as an example.[EB/OL].[2022.05.24]. https://dadh2020.ascdc.tw/#dadh-program.

[68] Jen-Jou Hung, Yu-Chun Wang, Shi Guanghui. Recommended method of famous buddhist sentences based on deep learning[EB/OL]. [2022.05.24]. http://dadh2021.ncue.edu.tw/pages/program.html.

[69] Shu-Jiun Chen, Hsiang-An Wang, Yu-Chien Ling, The international image interoperability framework (IIIF) approach to digital representation, view and annotation of buddhist caves and images, J. Digit. Arch. Digit. Humanit. 3 (2019) 63–100.

[70] Ming-Yeung Cheung, Yi-Fu Chen, Wei-De LI. The challenge of online religious language analysis: taking the online participation of christian religion in public issues in taiwan as an example[EB/OL]. [2022.05.24]. https://dadh2020.ascdc.tw/#dadh-program.

[71] Shih-Wen Chyu, Digital humanities for authorship attribution problem of shiji: based on frequency of function words within the hereditary houses section, J. Digit. Arch. Digit. Humanit. 2 (2018) 49–69.

[72] Ching-Syang Jack Yue, Yu-Ting Ye, A study of text mining on Taiwan's newspapers, J. Digit. Arch. Digit. Humanit. 6 (2020) 69–96.

[73] Shueh-Ying Liao, A quantitative research of the Book of odes "ShiJing": the discovery of the underlying rhythm in the incentive process, J. Digit. Arch. Digit. Humanit. 4 (2019) 49–70.

[74] Wenhui Zheng, Yu-Ting Ye, Yan jingxin, etc. The metaphor and metonymy of "He chu" A digital humanities study on the "poetics of space, in: Tang Poetry[C]. 9th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2018, pp. 379–383.

[75] Ma Lin, Monica Li, What helped officials of song dynasty in climbing the greasy Pole: an empirical study[C], in: 10th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2019, pp. 193–213.

[76] Yahui Xu, Tingshuo Huang, Huangyao Huang, Regional distribution and change of tomb cultural relics in South and north song dynasty: establishment and application of "Liao song Jin tomb database" [C], in: 9th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2018, pp. 270–287.

[77] Chiou-shing Yeh, Huei-ling Lai, Jyi-Shane Liu, The construction of taiwan Hakka corpus and preliminary analysis of Hakka lexical usage, J. Digit. Arch. Digit. Humanit. 8 (2021) 75–131.

[78] Ming Ren, Sherry Xie, Jin Zhang, Collective memory in the digital age: a case of beijing Drifters[C], in: 8th International Conference of Digital Archives and Digital Humanities Proceedings. Taiwan, 2017, pp. 18–21.

[79] Ping Wang, Liang Niu, Guan-Tao Jin, Etc. The digital humanities research of the landscape painting of the five Dynasties and northern song dynasty (2): a study of the "fisherman-hermit" theme in painting, J. Digit. Arch. Digit. Humanit. 1 (2018) 127–147.

[80] Wei-Min Fan, Muh-Chyun Tang, Creative role analysis of the best male/female Mandarin singer nominees in taiwan golden Melody Awards (2007–2018), J. Digit. Arch. Digit. Humanit. 3 (2019) 43–62.