

## NEWLY EXCAVATED TEXTS IN THE DIGITAL AGE: REFLECTIONS ON NEW RESOURCES

Lee-moi Pham\* and Kuan-yun Huang<sup>†</sup>

### Abstract

Based on a special issue entitled “Newly Excavated Texts in the Digital Age,” volume 21.2 (June 2011) of the Newsletter of the Institute of Chinese Literature and Philosophy, Academia Sinica (*Zhongyang yanjiu yuan Zhongguo wen zhe yanjiu suo tongxun* 中央研究院中國文哲研究所通訊), this article reflects on the various digital resources now being developed at institutions in China, Hong Kong, Taiwan and elsewhere, looking, in turn, at topics related to the graphic form of characters, the relation between character and word, and the question of context. In addition, the article considers the web forum, a platform of research and discussion that is increasingly becoming a part of scholarly exchange.

In the past few years, there have been several significant developments in the study of newly excavated texts with regard to digital and online resources.<sup>1</sup> In an effort to take account of this, we recently invited several scholars directly involved in these new endeavors to comment on their works. The results are published in a special issue entitled “Newly Excavated Texts in the Digital Age,” volume 21.2 of the

---

\* Lee-moi Pham, 范麗梅, Institute of Chinese Literature and Philosophy, Academia Sinica; email: [lmpham@gate.sinica.edu.tw](mailto:lmpham@gate.sinica.edu.tw).

<sup>†</sup> Kuan-yun Huang, 黃冠雲, Department of Chinese Literature, Tsing Hua University; email: [kuanyun@mx.nthu.edu.tw](mailto:kuanyun@mx.nthu.edu.tw).

1. For an early call to integrate digital technology with the study of newly excavated texts, see Qiu Xigui 裘錫圭, “Tuidong Gu wenzi fazhan de dangwu zhi ji” 推動古文字學發展的當務之急, in *Xueshu shi yu fangfa xue de shengsi: Zhongyang yanjiu yuan Lishi yuyan yanjiu suo qishi zhounian yantao hui lunwen ji* 學術史與方法學的省思: 中央研究院歷史語言研究所七十周年研討會論文集 (Taipei: Zhongyang yanjiu yuan Lishi yuyan yanjiu suo, 2000), 433–39. The growing importance of digital resources has resulted in serious discussions in some neighboring disciplines. An example of this is Huang Yi-long 黃一農, “Mingmo zhi Aomen mu Pubing de Jiang Yunlong xiaokao: jian da Xiong Xiong xiansheng dui ‘e-kaoku’ de piping” 明末至澳門募備兵的姜雲龍小考: 兼答熊先生對「e-考據」的批評, *Zhongyang yanjiu yuan Jindai shi yanjiu jikan* 中央研究院近代史研究所期刊 62 (2008): 155–66. The article includes references to earlier debates between Huang and other scholars on the significance and impact of “e-philology.”

Newsletter of the Institute of Chinese Literature and Philosophy, Academia Sinica (*Zhongyang yanjiu yuan Zhongguo wen zhe yanjiu suo tongxun* 中央研究院中國文哲研究所通訊) in June 2011.<sup>2</sup> Before we proceed to give our own take on this topic, it is perhaps helpful to give a summary of the articles.

Zhang Zaixing 張再興 (East China Normal University, Center for the Study and Application of Chinese Characters), “Establishing a Database of Newly Excavated Texts in the Age of the Internet: Reflections and Implementations.”

<http://www.wenzi.cn/web/content.aspx?moduleid=22&parentid=20>

This article lays out a principle of complete digitization, which involves the transcription of the original characters as well as their various components, also the use of tree diagrams to link the graphic variants. By assigning a code to each graphic form, a database operating with such a principle not only makes the graphic forms easy to input, but also enables the database itself to be fully searchable. In turn, the database provides the basis for additional databases specific to graphic form, lexicon and syntax. Such databases are already in operation, as with the databases on bronze inscriptions and the dating of bronze vessels, all created by East China Normal University.

Hung I-mei 洪一梅 (formerly of the Institute of History and Philology, Academia Sinica), “Developing an Intellectual Structure for the Study of Bone and Shell Inscriptions: Theory and Practice.”

This article drafts a developmental plan for an intellectual structure for the study of Shang bone and shell inscriptions. It gives a detailed account of two inter-linked databases: one featuring archaeological data ([http://ndweb.iis.sinica.edu.tw/archaeo2\\_public/System/Artifact/Frame\\_Search.htm](http://ndweb.iis.sinica.edu.tw/archaeo2_public/System/Artifact/Frame_Search.htm)) and the other the collection of bone and shell inscriptions at the Institute of History and Philology (<http://rub.ihp.sinica.edu.tw/~oracle/>). By laying out in detail the functions of the databases as well as their limitations, the article discusses the ongoing work on the database and the challenges ahead, and it also suggests various possibilities for improvement. Lastly, the article addresses the transformation of the study of bone and shell inscriptions in the new century based on social network theory and the idea of information sharing.

---

2. This as well as all previous issues of the newsletter are available on the website of the Institute of Chinese Literature and Philosophy (<http://www.litphil.sinica.edu.tw/home/6-2.htm>).

Ho Che Wah 何志華 (Research Centre for Chinese Ancient Texts, Chinese University of Hong Kong), "The Digitization of Newly Excavated Texts by the Research Centre for Chinese Ancient Texts, Chinese University of Hong Kong."

The Chinese Ancient Texts Database (<http://www.chant.org/>), with its long history and great popularity among scholars, now consists of three sub-programs related to newly excavated texts: one for bone and shell inscriptions, one for bronze inscriptions, and one for writings on bamboo and silk manuscripts. For each the project has developed highly adaptable search functions according to the particularities of the sources. Search results are accompanied by the transcription of the paleographic text for the convenience of users. In spite of its early success, currently there are plans to make the database more comprehensive, accurate, and user-friendly.

Cheng Shaoxuan 程少軒 (Center for Research on Chinese Excavated Classics and Paleography, Fudan University), "The Design and Production of the 'Comprehensive Database for the Characters and Words of the Shanghai Museum Bamboo Slips.'"

This article concerns the design of the "Comprehensive Database for the Characters and Words of the Shanghai Museum Bamboo Slips," with reflections on the lessons learned. The database is intended to retrieve relevant characters, vocabulary, and texts as well as secondary works pertaining to the Shanghai Museum manuscripts. It includes a program featuring individual characters and whole passages, including images of the original character, indices, transcribed forms, and suggested readings as well as the text files of the relevant secondary sources. In this way, four independent search functions are available: graphic form, semantic meaning, text, and secondary source. The article ends with the novel suggestion to open up the database and allow editing privileges to the scholarly community at large; this is to ensure that the database would be updated quickly to include the latest research.

Lu Jialiang 魯家亮 and Li Jing 李靜 (Center of Bamboo and Silk Manuscripts, Wuhan University), "The Study of Bamboo and Silk Manuscripts in the Digital Age."

<http://www.bsm.org.cn/>

This article discusses the work undertaken at the Center of Bamboo and Silk Manuscripts, including the design and application of an Internet interactive platform, database, and infra-red image technology. For the Internet interactive platform, the plan is to invite the participation of scholars from all over the world, providing a forum for new research,

the storage of scholarly resources, information sharing, among others. This helps to overcome the temporal and spatial boundaries of scholarly communication, expediting the study of newly excavated texts in both range and depth. In terms of managing the graphic forms and texts of the bamboo and silk manuscripts, the database developed by the Center enables users to retrieve information using either a single character or a graphic component, and the results provide the actual images of the characters, even the relevant portions of the manuscripts. Finally, the staff at the Center is now experienced in the use of infra-red image technology.

Shan Yuchen 畢育辰 (Institute of Ancient Documents, Jilin University)  
 “Newly Excavated Text and Related Database Resources in the Digital Age.”

This article surveys the various databases of bone and shell inscriptions, bronze inscriptions, silk and bamboo manuscripts, and transmitted literature. It also offers a number of suggestions from a user’s point of view as to the provision of digital images, single engine searches, and among others, the retrieval of Warring State scripts, whether as a whole or by region.

\* \* \*

As can be seen from the summaries above, a vast array of paleographic sources is now represented in digital form, ranging from the inscriptions on bones and shells to those on bronze vessels, seals, pottery vessels, coins, also writings on bamboo and wooden slips, jade and stone tablets, silk, among others. In some cases, a single database encompasses most if not *all* of the data from a single writing surface, as is the case with the bone and shell inscriptions of the Shang, based on the comprehensive *Jiagu wen heji* 甲骨文合集; or writings on the large number of bamboo and wooden slips from the Warring States, Qin, and Han. Further, there are plans for these databases to include additional sources relevant to the study of newly excavated texts. These include the *Shuowen jiezi* 說文解字, the Wei stone classics, and the *Hanjian* 汗簡 as well as the *Guwen sisheng yun* 古文四聲韻, two works compiled during the Song but now known to contain credible information transmitted from ancient sources.

As products of our times, these new digital resources are based on the assumptions, understandings, and consensuses currently held by the scholars in the field. But a question can also be raised—to what extent do these resources help to advance, modify, or even radically change the discipline? What follow are our reflections on the digitization of newly excavated texts, offered in the hope of highlighting some of the

most prominent features of the various resources currently being developed. The discussion is divided into two parts. The first comments on topics related to the graphic form of characters, the relation between character and word, and the question of context, while a second part considers a platform of research and discussion that is increasingly becoming a part of scholarly exchange: the web forum. We end with some thoughts about the future of scholarship.<sup>3</sup>

### Graphic Form of Characters

For the graphic form of characters in paleographic texts, an important question is how can the rich amount of information embedded in the graphic form of a character be accurately reflected and conveniently accessed in digital form? Here, by “graphic form,” we refer to the calligraphic tendencies of a character, the composition of the strokes, and the positioning of the graphic components, generally called the *pianpang* 偏旁. Such information, to be sure, is available from the photographic reproduction, but when it comes to a digital database, the question is how this information should be inputted and extracted. In many ways, this is a question that overlaps with the matter of the transcription of paleographic texts. The most important issue at stake is what is the standard adopted when a paleographic character is being inputted into a database? As for the transcription by the original editors, should this be retained, for the sake of consistency and the completeness of the record, even when there is clear evidence that it is imprecise or even mistaken? Or should one try to come up with a new transcription that then serves as the basis of the database? To look into the matter a bit more deeply, the transcription of a character is often inseparable from one’s understanding of the variations of a character, in both diachronic and synchronic (e.g. trans-regional) terms; it is questionable whether it can be resolved in the single-step process that is the transcription, let alone dictate the production of a database. Such an understanding

---

3. Some popular books that have been helpful in our thinking through some of the issues addressed in the following are James Gleick, *The Information: A History, a Theory, a Flood* (New York: Pantheon Books, 2011); David Weinberger, *Everything Is Miscellaneous: The Power of the New Digital Disorder* (New York: Times Books, 2007); and Alex Wright, *Glut: Mastering Information through the Ages* (Washington, D.C.: Joseph Henry Press, 2007). The following reference works contain a wealth of information: Susan Schreibman, Ray Siemens, and John Unsworth, eds., *A Companion to Digital Humanities* (Oxford: Blackwell, 2004); and Ray Siemens and Susan Schreibman, eds., *A Companion to Digital Literary Studies* (Oxford: Blackwell, 2007). A useful account of what “digital humanities” is (or is not) can be found in Anne Burdick, Johanna Drucker, Peter Lunenfeld, Todd Presner, and Jeffrey Schnapp, *Digital Humanities* (Cambridge: MIT Press, 2012).

will ultimately lead a serious scholar back to the beginning of his or her reflection: newly excavated texts are by and large manuscripts, written by hand, and this is the feature that distinguishes them from the printed texts that we have become accustomed to.

In response to these questions, the various institutions strive for a method for conducting multi-level searches, and also to gradually develop a unified standard for the transcription. In terms of multi-level searches, Wuhan University has led the way in offering the feature of searching by graphic components. As noted by the scholars at that institution, this feature is especially crucial for those cases where one recognizes only one part of a character and not the rest. By accessing in a systematic manner all the characters featuring a particular graphic component, paleographers can identify whole series of characters, sometimes phonetic series. To address the same problem, the database of the manuscripts from the Shanghai Museum, created by Fudan University, has an index that itemizes every character and all the components. Such features can also be found in the database developed by East China Normal University.<sup>4</sup>

In terms of setting a unified standard for the transcription of characters, Academia Sinica has for some time pushed for the creation of a database of character components, or *Hanzi gouxing ziliao ku* 漢字構形資料庫, and its database of bone and shell inscriptions shows the correspondence between the inscriptions and transcribed characters. In contrast, East China Normal University, with its principle of complete digitization, assigns a code to every character, which can then be inputted using a method devised by the same institution. Finally, Fudan University has a new idea for how to search for secondary sources, namely, by employing a unified and standardized transcription prepared by the architects. Together, these efforts raise the question of whether a single, practical, and plausible standard could be developed by the architects of these databases, so ultimately there would be a unified list of characters as well as an input method that transcends institutional boundaries.<sup>5</sup>

---

4. Here it is worth noting a project developed by Kwan Tzu-wan 關子尹 and his team from the Department of Philosophy at the Chinese University of Hong Kong, which analyzes (*à la* the *Shuowen jiezi*) characters into minimum components and organizes such components into a hierarchical structure. A part of this project is described in Kwan's proposal for the project, "Multi-function Chinese Character Database: An Infra-structure for Future Chinese Education" (<http://qcrq.qef.org.hk/project.php?id=2009/0456>).

5. Such an effort is now reportedly under way. The project "Zhonghua ziku" 中華字庫, led by Qiu Xigui of Fudan University, aims to create a comprehensive database of Chinese characters ([http://www.gwz.fudan.edu.cn/srcShow\\_NewsStyle.asp?Src\\_ID=1612](http://www.gwz.fudan.edu.cn/srcShow_NewsStyle.asp?Src_ID=1612)).

Finally, it is possible to mention a feature developed by the Chinese University of Hong Kong, an input method by handwriting. Imagine if one wrote the paleographic form of a character, and upon input, received a series of characters that resemble one another, however “resemblance” is defined. The result would not only provide information relevant to the problem of decipherment at hand, but also draw attention to other aspects critical for the recognition of a character. To be more specific, if characters that resemble each other in terms of the calligraphy, stroke, and component could be digitally grouped in a particular manner, then such series of characters with similar or identical variation patterns could provide new ways for understanding the development of the writing system in general.<sup>6</sup>

### Relation between Word and Character

How do the characters in paleographic texts record the spoken words of Old Chinese? This is the most fundamental as well as the most difficult question in the study of newly excavated texts. In its database of the Shanghai Museum manuscripts, Fudan University provides a table indicating the correspondence of characters and words, or *Zici duiying shuju biao* 字詞對應數據表, and by allowing multi-functional searches, makes possible the building of a network of closely related characters and words.<sup>7</sup> In connection with this, the Fudan database indicates for each character the Old Chinese initial and rhyme group, following the reconstructions by Wang Li 王力 (1900–1986), and this is also a feature of the database created by East China Normal

---

6. Related to this, it is possible to mention the *xizi* 習字 or practice exercises frequently seen among Han administrative documents from the northwest. These writings, in spite of their clumsy and unskilled execution, pose interesting challenges to the current understanding of the standards and procedures for transcribing a character. In addition, there are various kinds of non-linguistic symbols appearing on newly excavated texts, conventionally referred to as *mogou* 墨鉤 “ink hook,” *moding* 墨丁 “ink block,” among others. These items are useful for understanding the layouts of manuscripts, their internal organizations, among others, but they are generally not included in databases.

7. This is also a feature attempted in both the *Xian Qin jiagu jinwen jiandu cihui ziliao ku* 先秦甲骨金文簡牘詞彙資料庫 developed by Academia Sinica and the project by Kwan Tzu-wan. In both cases, under each character, one finds a list of words that the character corresponds to. Kwan’s project is especially noteworthy because it also lists words that are homophonous, thus facilitating the search for phonetic loans. In addition, each entry contains an explanation of how the different meanings relate to one another.

University.<sup>8</sup> It is not difficult to see that if closely or identically sounding words could be obtained through database searches, this would lead to unexpected results in the decipherment of a character. Related to this, one might raise the more daring question of whether it might ever be possible to search by semantic meaning, so words with close meanings could be linked?<sup>9</sup> If nothing else, this would allow one to observe phonological connections that might otherwise go unnoticed.

When considering the question of the relation between word and character, it is important to note that a text sometimes has parallels with other texts, whether this means that it is quoted by another text or that two texts, to whatever extent, have overlapping content, vocabulary, or syntactic form. These correspondences are valuable sources of textual variants, allowing for comparisons on several levels. Touching on the complex issue of how to compare newly excavated texts and the transmitted literary record, these parallels are a reminder of how texts can vary across time and space and from one transmitter to another, thus providing an important resource for the discussion of textual transmission. It is not inconceivable that if such data were linked digitally, then the relevant information could be made searchable

---

8. The study of Old Chinese phonology has seen several new developments recently; especially noteworthy are the hypotheses presented by the comparative study of Old Chinese with languages from the Tibeto-Burman and even Austronesian families. Besides the ongoing work by William Baxter and Laurent Sagart, it is possible to mention Zheng Zhang Shangfang 鄭張尚芳, *Shangwu yinxi* 上古音系 (Shanghai: Shanghai jiaoyu, 2003), and Axel Schuessler, *ABC Etymological Dictionary of Old Chinese* (Honolulu: University of Hawai'i Press, 2007), both with reconstructions that are easy to look up. Further, there now exist several online resources. The first is the website <http://www.eastling.org/>, which features the reconstructions by Bernhard Karlgren, Li Fang-kui 李方桂, Wang Li 王力, Baxter, Zheng Zhang Shangfang, and Pan Wuyun 潘悟雲. More recently, under the direction of the Department of Chinese Literature at National Taiwan University and several institutes at Academia Sinica, a new website is launched, <http://xiaoxue.iis.sinica.edu.tw/ccr/>, which provides Old Chinese reconstructions by several scholars, among other valuable information. These resources would no doubt play a significant role if the aforementioned databases of newly excavated texts were ever to expand in the direction of Old Chinese phonology.

9. This is attempted in the project directed by Kwan Tzu-wan, which divides various components of characters into seven major categories: the human body, bodily gestures and physical behavior, natural phenomena, plants and animals, life and culture, deictic (*zhishi* 指事) characters, and those with multiple or uncertain meanings. In addition, it identifies the word class as well as provides an English translation. But the database with the most detailed treatment of these matters is perhaps the Thesaurus Linguae Sericae (TLS) created by Christoph Harbsmeier, University of Oslo ([http://tls.uni-hd.de/home\\_en.lasso](http://tls.uni-hd.de/home_en.lasso)). Among the rich information found under each entry are phonological profile, synonym groups, synonym groups definition, syntactic categories, and semantic categories.



in an effective and comprehensive manner. On this point, the database by the Chinese University of Hong Kong provides the following feature: for a newly excavated text, if there is a relevant transmitted source, then this is cited and compared, resulting in emendations (indicated by a different color font), critical notes, or tables summarizing the differences. In addition, the Chinese University has published several reference works that compile these quoted and cited sources with respect to particular ancient texts, in a series entitled the *Xian Qin Liang Han dianji yin jingdian ziliao huibian* 先秦兩漢典籍引經典資料彙編. On this basis, one wonders if statistical methods or other tools could be utilized to come up with a more precise definition for parallel texts, which would in turn allow for the identification of additional relevant materials?

### Context

The paleographic research of the past has generally focused on the decipherment of single characters. The type of reference work that best reflects this interest is the index of graphic form, or *wenzi bian* 文字編 (and the concordances that are comparatively fewer in number). Due to the growing variety and complexity of newly excavated texts, particularly literary and philosophical texts, scholars sometimes propose readings and interpretations that vary widely from one another. The context for placing these paleographic sources is increasingly becoming a challenge, and it has a direct impact on using the new evidence to make certain claims about language, thought, culture, and other aspects of Early China. Here “context” can refer to a bamboo slip, a whole text, various documents from the hand of a single scribe, a single tomb, or it can be a web or interpretative framework based on a particular researcher’s understanding of language, thought, culture, and other historical backgrounds. How should these different “contexts” be reflected so they could meet the needs of scholars—this is a question that has no ready answer.

Through their databases, different institutions respond to this question in different ways. One response is to link a search result with the photograph or rubbing of the corresponding object, thus reproducing the context as it is reflected in that impression. This is the case for the Chinese University of Hong Kong in its database of the bone and shell inscriptions, bronze vessel inscriptions, and bamboo slips and silk. It is also true for the series of databases designed by Academia Sinica, which include not only the database of bone and shell inscriptions under discussion, but also a database for bronze inscriptions of the Shang and Zhou. In its database, East China Normal University divides the various levels of transcription and suggested reading, and

it takes into account the correspondences as well as possible discrepancies between the paleographic text and later decipherments. It has also published a book series that gathers the various opinions of paleographic research, character by character, called the *Gu wenzi kaoshi tiyao zonglan* 古文字考釋提要總覽, with a corresponding searchable online database. Finally, Fudan University has a bibliography of secondary studies that links the relevant scholarship to a corresponding character of a paleographic text; and in its transcription and reading, it aims to incorporate the best and latest research.

### Web Forums and Publication

Given the frantic pace at which newly excavated texts are coming to light, it is little surprise that more and more scholars are opting for web forums as opposed to print publication for disseminating their research. Rather than waiting for months or years for an article to come into print, one can share one's ideas almost instantaneously. Among the contents featured on these forums are research notes; previously published papers; announcements of projects, conferences, and other scholarly events; and publication notices. To name only some of the most active web forums:

- Fudan University, Center for Research on Chinese Excavated Classics and Paleography (<http://www.gwz.fudan.edu.cn/>)
- Wuhan University "Bamboo and Silk Manuscripts" (<http://www.bsm.org.cn/>)
- Tsinghua University, Unearthed Research and Protection Center (<http://www.tsinghua.edu.cn/publish/cetrp/>)
- Chinese Academy of Social Sciences, Institute of History (<http://www.xianqin.org/>)
- "Bamboo Silk" (<http://www.jianbo.org/>)
- "Confucius 2000" (<http://www.confucius2000.com/>)

At the forum hosted by Fudan University, it is not uncommon for a research note to be followed by readers' comments, rejoinders, and sur-rejoinders. This is not the case for every forum, though there is a great deal of cross-referencing among the research notes, an indication of the high level of frequency at which they are consulted. While some of the forums regularly bundle together their contents—both Wuhan University and Fudan University gather the research notes and make them available as compact discs or electronic files—a question that

naturally arises is whether the forums would in the future introduce a peer-review process, which is ultimately what separates these online discussions from print publications. (For some of the institutions that already publish printed journals, the question can be put somewhat differently, as to whether they would integrate the forums with their journals.)<sup>10</sup> As long as the refereed articles remained distinct from other discussions taking place online, there is no risk of their conflicting with one another. The advantages would of course be accessibility for all, the saving of production costs, and the possibility that an article, once published, could provide the basis for further debates and discussions, a living document, so to speak.<sup>11</sup>

The major question with the web forums as they currently operate is of course their impact on scholarship. While the discussions often give the impression of diversity, perhaps even cacophony, they also have the paradoxical effect of impeding truly innovative ideas. Just as one idea is quickly echoed and variously referenced so it assumes the appearance of accepted consensus, another idea is submerged and forgotten, never receiving the consideration that it rightly or not rightly deserves. In this regard, what the sociologist James A. Evans suggests about online journals could just as well be true for web forums. Basing his research on a database of 34 million articles, their citations, and online availability, Evans argues that an increase in online availability actually has the adverse effect such that "the articles referenced tended to be more recent, fewer journals and articles were cited, and more of those citations were to fewer journals and articles." According to Evans, "Searching online is more efficient and following hyperlinks quickly puts researchers in touch with prevailing opinion, but this may accelerate consensus and narrow the range of findings and ideas built upon."<sup>12</sup> While the situation for the study of newly excavated texts is somewhat unusual, with the web forums, journals, and the large number of electronic books circulating on the Internet occupying different levels of scholarly discourse, it is not immune from the problems described by Evans, and the implications are well worth pondering.

---

10. Some papers from the "Bamboo Silk" website have been reissued in print form; see *Gumu xinzhī* 古墓新知 (Taipei: Taiwan guji, 2002) and *Guodian Chujiang yu zaoqi Ruxue* 郭店楚簡與早期儒學 (Taipei: Taiwan guji, 2002).

11. For some of the same reasons, one might push for more academic conferences to take place via the web.

12. James A. Evans, "Electronic Publication and the Narrowing of Science and Scholarship," *Science*, July 18, 2008, 395–99.

### Potentials of the “Wiki” Approach

The success of Wikipedia and its popularity among students and the general audience have led many to consider more seriously the possibility of an open, collaborative approach to scholarship, particularly the preparation of such reference works as the encyclopedia.<sup>13</sup> Is this approach possible for the study of newly excavated texts? Although it is understandable that the architect of a database might not wish to open his or her work to just about everyone—Cheng Shaoxuan of Fudan University mentions the possibility of a vouching process that would allow certain scholars the privilege of modifying the database developed by Fudan—there are more fruitful areas where a collaborative approach might be attempted. One possibility is the compilation of bibliographies. As noted in the preceding section, publication has now reached such a frantic pace that original research is published daily as monographs, journal articles, chapters in edited volumes, and research notes at web forums (including the readers’ comments) all clamor for one’s attention. It is difficult for one person or even a group of scholars to keep track of all these, let alone read everything; and the problem is exacerbated as one moves beyond one’s native language. A collaborative approach has the potential to solve many of these problems; in fact, it might be the only practical solution. With regard to the problem of vandalism that is often mentioned as a potential risk of collaborative approach, this has already been discussed in connection with Wikipedia and can easily be overcome by such simple measures as requiring registration only by one’s real name and maintaining a certain level of administrative oversight.

Once a bibliography has been established, it can be organized according to the various principles of library and information science, and retrieved by searchable options, including keywords and subject headings. At the same time, from the user’s perspective, if one could have the option of attaching a certain label (in the form of keywords or short phrases) to individual bibliographical items, a practice known as “tagging,” then this might open up even greater possibilities, one of which would be to reveal all kinds of interesting trends about how scholars receive and respond to the works of their peers.

A collaborative approach has great potential also in terms of pedagogy. One can envision such an approach for course syllabi, lecture notes, textbooks, and other areas where consensus and mainstream opinions are

---

13. For an useful discussion of Wikipedia and the challenges that it presents, see Roy Rosenzweig, “Can History Be Open Source? Wikipedia and the Future of the Past,” *Journal of American History* 93.1 (2006): 117–46.

valued over individual innovation. While beginning students are not expected to know the full range of opinions on a certain question, they should have a sense of the most important and most historically interesting. This is where the input of others can be beneficial, and the result, as the product of collective labor, would demystify a field which many outsiders regard as arcane, out of date, and tirelessly quarrelsome, and present it in a more favorable light. Note that such an approach is fundamentally different from the designing of websites, which, like museum exhibitions, can reach and touch the general audience in ways unimaginable by the printed word. Commendable as such projects are (and many do make the effort to be balanced and even interactive), their approach is ultimately a top-down, one-directional creative process.

Come to think of it, why would the Wiki approach not work for databases? It is without question that a database, like all other reference works and systems of knowledge, requires structure. But if such structure could be laid down in a way that permits the participation of its users, then this would only benefit the field as a whole. Thus, for a problematic character in a newly excavated text, it is no longer one or several scholars who are responsible for noting the large number of readings that have been proposed; everyone can chip in. As long as there is a mechanism to distinguish the different opinions, so the view of the original editors is differentiated from that of the architect, and in turn from the other proposals, then the work avoids the risk of alienating any of its users. The result would resemble the type of reference work, the *gulin* 詁林 or *jishi* 集釋 that scholars have become accustomed to for the study of paleographic texts, but it would be an ever-expanding one, with an endless list of addenda already incorporated into the database itself.<sup>14</sup>

### Future Prospects

For the present discussion as well as the study of newly excavated texts in general, the focus is and always will be on the reading of texts. While the written record of language cannot be examined in isolation from archaeology, art history, and the materiality of the writing surfaces, among other concerns, the priorities are beyond questioning. For a time when new approaches and methodologies are constantly being advanced and debated, with the field flooded by an ever-expanding body of data, we believe this point is worth emphasizing. As one looks on to a new age, one must never forget that the digital resources

---

14. This is of course also true for databases of transmitted texts, for which commentaries are found not only in the standard editions, but also in various formats in the scholarly literature. Once again, a collaborative approach can be helpful.

are intended to reinforce and supplement the traditional approaches, never to replace them.

Here it is helpful to look to history and recall that the reproduction, management, and processing of ancient textual sources has been an ongoing project, traceable, at least in the modern period, to the Sinological Index Series published under the direction of William Hung (Hong Ye 洪業, 1893–1980). From 1931 to 1950, Hung and his colleagues at the Harvard-Yenching Institute engaged in the itemization of the linguistic data of seventy-seven ancient works, basing their work on the critical texts that they had established and making the efficient and comprehensive use of these sources a reality.<sup>15</sup> Such endeavors, as Hung explains in his essay “On Indexing” (*Yinde shuo* 引得說), were not unprecedented and followed in the footsteps of the scholars of a previous generation.<sup>16</sup> In fact, if one included dictionaries, rime books, encyclopedias, bibliographies, and various guides and reference works, all of which reflect the diverse ways in which traditional scholarship confronted the seemingly boundless ocean that is the transmitted literary record, then the effort to effectively manage large quantities of textual information may go back even farther, and it can perhaps be understood as a recurring theme in the possession, preservation, editing, sorting and even censoring of texts. Such is the paradox that in attempting to manage information, we end up creating more of it. Such creations, like all textual creations, say as much about a literary tradition as they do about ourselves.

So where do we go from here?

There are, to be sure, many problems and challenges presented by the new resources, and some of these, such as the “narrowing” of scholarship, we alluded to above. In our view, what is most significant about these new developments is the following. If, in the past, we have relied on authorities and standard accounts, we will continue to do so, for that is how we acquire knowledge and make sense of the world. But the new resources also suggest a conception of scholarship and much else that is fundamentally different from what we are used to. Rather than a field dominated by a handful of towering figures, we now have multiple voices; rather than one character, one reading, we now have a dozen proposals, maybe more, for how the text should be understood; rather than a stable corpus, we now have an ever-expanding body of data; rather than a fixed standard text, we now have cross-

---

15. Under the direction of the late D.C. Lau (Liu Dianjue 劉殿爵, 1921–2010), the Ancient Chinese Texts Concordance Series of the Institute of Chinese Studies, Chinese University of Hong Kong, is a direct continuation of Hung’s work.

16. Hong Ye 洪業, *Yinde shuo* 引得說 (Beiping: Hafo Yanjing xueshe yinde bianzuan chu, 1932).

referencing links pointing to widely diverging interests and agenda; rather than a center of research based in Beijing, Paris, Kyoto, or Berkeley, we now have scholars, graduate students, and amateurs working from all over the world. As much as old habits will endure and traditional approaches continue to be upheld—and we emphasize again that we do not advocate that they should be abandoned—it seems that these efforts now take place within a context that is ultimately diverse, multi-polar, democratic, and non-hierarchical, with the participation of the whole scholarly community. Any claim for authority and final answer will be as much a textual creation as the databases and other resources described here. Instead of asking what the correct reading of a character is, perhaps we could also ask what are the conditions that give rise to the multiplicity of meaning, the rationale underlying each proposal, the personalities behind them, and the various historical contexts in which they are embedded. If this implies anarchy, it also means that we will look more carefully at differences, potentials, and the spaces between the lines. The mode of inquiry that results is more open-ended, and scholarship becomes a process whereby each person, as an individual, makes his or her contribution to the community at large. For a discipline so steeped in tradition, and in a world overshadowed by authoritarianism in various forms and guises, such an understanding, as with open access, freedom of speech, and other related causes, is ultimately to empower the individual and ensure a more humane way of life, and it is worth striving for.

## 數位時代的出土文獻：關於幾種新工具的反思

范麗梅、黃冠雲

### 提要

本文根據《中央研究院中國文哲研究所通訊》2011年出版的第21卷第2期，「數位時代的出土文獻」專輯，討論近幾十年來在大陸、香港、台灣等地區，由不同單位開發的各種數位工具，從使用者的角度對字形、字詞對應、語境等幾個方面提出感想。此外，論文也考慮互聯網論壇在研究和交流上所扮演的日漸重要的角色。

**Keywords:** newly excavated texts, paleography, digital resources, web forums, Wiki approach

出土文獻, 古文字, 數位資源, 互聯網論壇, 「維基」模式