

The Digitized Tripitaka Koreana 2004: Benefits and Challenges in East Asian Buddhist Studies¹

Kim Jongmyung

The purpose of this paper is to examine the history of the digitization of sources for Korean studies, benefits from the Digital *Tripitaka Koreana 2004* (*DTK 2004*), the newest digitized version of the *Goryeo daejanggyeong* (Tripitaka aka Koreana, *TK*), challenges to it, and future directions for it with regard to the study of East Asian Buddhism. This paper approaches from the perspective of an end user of the *DTK 2004*, without any special claims of technological expertise. The *DTK 2004* hopes contribute to promoting Buddhist studies throughout the world. However, it has both benefits and challenges. Its strengths include the revival of the original text, more useful functions such as comparison between the *TK* and the *Taishō shinshū daijōkyō*, and its potential contribution to the cultural studies of East Asia. It also faces some challenges, including the lack of a printed manual and photo-images of the woodblock prints of the *TK*, limited access to browsers, no links or mark-ups, the use of mixed romanization, and no further editorial effort. However, if we want better digital resources and a more responsible use of them, we must first resolve these problems. Furthermore, issues of participation of a wide range of specialists, the establishment of better planning, the setting up of standardization, the maintenance of high quality, and the economics of the *DTK 2004* remained to

-
1. The revised version of Kim Jongmyung 2005, this work was supported by the 2005 Academy of Korean Studies grant. In particular, I owe much to Venerable Daeseok, Jeon Cheolhwan, a program developer, and others of the Research Institute of the Tripitaka Koreana (RITK) for this research. I have had interviews with them in regard to the characteristics, functions, merits, and demerits of the *DTK 2004*. I also appreciate valuable comments from Professors Heo Hungsik and Lee Jongcheol of The Academy of Korean Studies.

be resolved. In particular, the majority of challenges to the further usability of the *DTK 2004* are primarily concerned with people in charge and cost. Buddhist specialists, computer scientists, and other related scholars should cooperate for the lasting editorial works of the *DTK 2004*. The *TK* is a National Treasure of Korea and the Depositories for the *TK* (*Janggyeonggak*) are now registered on the UNESCO world heritage list. In addition, fundraising for the constant maintenance and technical advancement of the *DTK 2004* is the most urgent impending issue. Therefore, government funding is of prime importance for the ongoing preservation and accessibility of the *DTK 2004*.

Keywords: Buddhist studies, digitization, East Asia, *Tripitaka Koreana*

Introduction

Religious groups have enthusiastically embraced the digital era. In particular, many countries have digitized various editions of Buddhist canonical texts. In recent years the availability of digital primary sources and other digitized resources has had an ever-increasing impact on Asian Buddhist studies. New digital materials are constantly appearing and more and more scholars are turning to these resources. There is almost universal agreement that the use of digitized source materials can enhance scholarship. Digitization provides researchers with easy access to large amounts of analysis or comparison of texts and data with potent search capabilities. This makes textual research more effective and convenient. It also enables scholars to engage in comprehensive data gathering and analysis, or comparative philological research. The easier preservation of data, the use of various types of data, including photos, sound, and motion images, the advancement of research productivity, and the portability of computer technology (Kim 1988: 296-8; Choe 2001: 399-400; Park 2004: 263; Poceski 2005) are also benefits available from the digitization of data. However, the problems that might arise with their use remain largely unexamined.² For exam-

2. These problems were extensively discussed in the panel, "Windfalls and Pitfalls: The Use of Digitized Material in the Study of Asian Buddhism," at the Association for Asian Studies (AAS) annual meeting held in Chicago, U.S.A., on March 31-April 3, 2005.

ple, production and publication of digitized sources have usually not been subject to the review and quality control process that scholarly publications normally go through (Schlutter 2005).

The Digital *Tripitaka Koreana 2004* (hereafter, *DTK 2004*),³ the newest digitized version of the *Goryeo daejanggyeong* (Tripitaka Koreana, hereafter, *TK*), is comprised of three CD-ROMs: DISC 1 *Geomsaek enjin* (Search Engine, *DTK 2004_1*), DISC 2 *Yeongu jeongbo* (Research Information, *DTK 2004_2*), and DISC 3 (Supporter CD). The *TK* was not the first effort of its kind in this area. However, it was the first one that survived intact down to modernity. The *TK* remained as the leading authoritative source for the East Asian Buddhist canon until its usage outside of Korea was surpassed by the Japanese *Taishō shinshū daijō kyo* (hereafter, *T*) (Muller 2000: 39), a collection of Buddhist scriptures which was compiled between 1921 and 1934 in Japan. The purpose of this paper is to examine the history of digitization of sources for Korean studies, benefits from the *DTK 2004*, challenges to it, and future directions for it with regard to the study of East Asian Buddhism. This paper approaches from the perspective of an end user of the *DTK 2004*, without any special claims of technological expertise.⁴ Among the three CD-ROM set of the *DTK 2004*, DISC 3 contains the names of supporters for the publication of the *DTK 2004*. Therefore, this paper is only concerned with DISC 1 and DISC 2.

History of the Digitization of Classics in Korea

The year 1981 was the first year of the era of the digitization of Korean studies (Kim 2005: 6). The digitization process of sources for the discipline has been particularly carried out in the field of history and Buddhist studies. The digitization project of sources of Korea was pioneered by Professor Wagner of Harvard University (Choi 2001: 386). He attempted to digitize the comprehensive rosters for the higher civil service examination in the Joseon dynasty (1392-1910) in the 1960s, which was called “the *Munkwa* Project.” The purpose of this project was to describe the computer-assisted *Munkwa* Project and to look at the problems

3. For the history, structure, and contents of the *TK*, see Kim 2002 (pp. 154-9). For the content of the *DTK 2004*, refer to Kim Jongmyung 2005.

4. Such an approach was also undertaken by Wagner and Poceski.

and the potential of this kind of undertaking in Korean historical research (Wagner 1982: 144).

Indebted to Wagner, the first official discussion of the digitization of sources for Korean studies in Korea was at the international conference titled, “Hangukhak jaryoui jeonsanhwa yeongu” (A Study on the Digitization of Sources for Korean Studies), hosted by The Academy of Korean Studies in Seongnam, Korea in 1981 (Kim 1994: 3). In addition, the first digitization project of historical sources of Korea was conducted under the sponsorship of The Academy of Korean Studies in 1984 (Choi 2001: 389). The field of history in Korea has made a strenuous effort to digitize research materials, including *Samguk sagi* (History of the Three Kingdoms), *Samguk yusa* (Memorabilia of the Three Kingdoms), *Goryeosa* (History of the Goryeo Dynasty), and the *Joseon wangjo sillok* (Veritable Records of the Joseon Dynasty) (Yi 2000: 154). In particular, the digitization of the *Veritable Records of the Joseon Dynasty* in 1995 was an epoch-making project of its kind in Korea in that it expedited the digitization of historical sources of Korea (Choe 2001: 385-9) and served as real information for research (Yi 2000: 158). Subsequent studies also examined directions for the digitization of data on Korean studies. Thus far, the digital historical sources exceed 60 (Park 2004: 259-89).

In Buddhist studies, a complete Chinese translation of Buddhist literature was first digitized in Korea based on the *TK*. As one of the major achievements of East Asian Buddhism, the *TK* is the Korean edition of the Chinese canon of Buddhist literature created on more than 80,000 wooden printing blocks in the thirteenth century, and presently stored at the Haein Monastery in the Republic of Korea. The *TK* is now available to readers in a 48-volume set of facsimiles of the xylograph rubbings.⁵ Published by Dongguk University Press in Seoul in 1976, this photolithographic reprint edition allows easy access for a large number of readers.

With the advent of the information society, the RITK of the Haein Monastery in South Korea undertook the digital input of the *TK* in revisions. The first edition was the Samsung edition, which was created on Humin Jeongeum software in 1993. This creation was historic in that it was the first digital edition of the *TK* in its entirety. However, it did not include a professional search function. The second edition was created in 1997 using the 4-byte code based Goryeo Word

5. The copper edition of the *TK* is now under consideration in contemporary Korea.

software. This was to manage the overwhelming number of variant forms (*iche-ja*)⁶ that were discovered. However, this again appeared to be a hindrance when searching the digital *TK*. Therefore, it was decided to include the Unicode edition with the *DTK 2000* edition (Hur 2000: 21-3) supported by the Ministry of Culture and Tourism. This was the world's first CD-ROM containing a complete set of Buddhist texts in classical Chinese. However, due to some errors and the difficulty of using the various functions, the result of this effort was not satisfactory. Based on the digitized version of the *T* by the Chinese Buddhist Electronic Text Association (hereafter, CBETA)⁷ and sponsored by the Haein Monastery, the revised version, the *DTK 2001*, came out in 2001. However, the *DTK 2001* still had the problem of variant characters. Therefore, the members of the RITK set out to develop a further edition,⁸ which resulted in the *DTK 2004*, the newest of its kind.⁹

Benefits from the *DTK 2004*

The texts compiled in the *TK* Woodblocks played a pivotal role in the development of the East Asian canon (Muller 2004: 3). It is also said that we need to study Korean Buddhism to better understand the development of pan-Asiatic Buddhism (Buswell 1989: xiv). In such a context, the historical significance of the electronic input of the *TK* is immense, in that a complete set of Chinese translations of Buddhist literature has been digitized for the first time, thus enabling students of East Asian Buddhism to gain easier access to Buddhist scripture (Kim 2002: 172).

The *DTK 2004* has the following characteristics. It is an integrated text retrieval system that enables one to search for and view *sūtras* and refer to annotation, the *Dictionary of Buddhist Terms*, and the *Electronic Chinese Character Dictionary*. One also can view the original text of the *TK* and the *T*, as well as

6. *Icheja* means the characters that are distinguished from standard characters according to slight visible differences in individual strokes and character constitution and structure (Lee 2000: 91).

7. With regard to this, Venerable Jongnim, Director of the RITK, expressed his deep gratitude to the CBETA (RIDTK 2004: *hugi* [postface]).

8. For the prime points in the 3rd correction, refer to The Prime Object on DISC 1.

9. It is said that more than 100 specialists have participated in the digitization of the *TK* for over ten years and over 80 billion Korean Won (approximately US\$8 million) were spent for the project.

images of the photo-prints, and one can use basic functions such as text editing, copying, and printing. One also can print out the photo-prints. The Chinese character input editor automatically shows the original text from the *T* that matches the passage of the *TK* that one is reading. As such, one can easily compare the two editions. For user convenience, the program also automatically provides a distributional analysis of characters and the *Handy Chinese Character Dictionary*. In particular, the revival of the original text, with more useful functions than its previous editions, and the potential contribution to the cultural sciences are the most important merits of the *DTK 2004* in the study of East Asian Buddhism.

1. The Revival of the Original Text

Most important in the process of the digitization of literary collections is entering the original data (Gwon 2004: 230). The *DTK 2004* was developed under the principle of the revival of the original text, engraved in the *TK*.¹⁰ However, the greatest hindrance to the digitization of sources in classical Chinese was the treatment of Chinese characters (Kim 1988: 304-10; Choe 1994: 88), including the limited number available in present computer software. In addition, due to the existence of deformed or variant characters in the *TK*, there was inaccuracy in the content. Therefore, under the principle of faithful representation, the members of the RITK made special fonts for the Chinese characters that cannot be represented by computer and then were able to produce the computerized *TK*. The types of variant characters and the total number found in the *TK* are 29,475 and 14.5 million, respectively (Lee 2000: 91). The types of variant characters found in the *TK* are as follows: The addition or subtraction of individual strokes; alteration of individual strokes; addition or subtraction of character components; altered placement of the character components; alteration or discarding of a structural component; and others (Lee 2000: 91). The identification of variant characters in the process of digitization of the *TK* was important, and the publication of the *Goryeo daejanggyeong ichejajeon* (Dictionary of Variant Characters in the *TK*), which contains 29,478 kinds of variant characters found on 1,477 pages, was the product of this process. In addition, the font for express-

10. The principle of revival in the process of digitization was to input the original texts as they were, although it proved to be an error.

ing the variants, which was created in the process of the digitization of the *TK*, is comprised of over 70,000 characters, and there are now 30,000 characters in use (RITK 2000: 3).

The *DTK 2004* is expected to be more easily accessible to end users in the field of East Asian Buddhism. In particular, the list of variant characters and the creation of the font for expressing them will contribute to solving technical problems in digitizing other canonical texts in Sanskrit, Pāli, and Tibetan.¹¹ In addition, the experience of the electronic input of the *TK* also will help solve financial problems encountered in digitizing these editions of Buddhist texts.

2. More Useful Functions

The *DTK 2004* has more useful functions than its previous versions, including diverse research information such as *sūtras* search, *sūtras* annotation, electronic dictionaries, character search, text editing, copying, printing, and comparison between the *TK* and the *T*. One can conduct an overall search for *sūtras* or view *sūtras* by searching the *sūtras* title or the *sūtras* number. One can view the *sūtras* of the *TK* and the *T*, and the titles and numbers of the *sūtras* are provided. One also can select several *sūtras* and view them all at once. Through the full-text search for *sūtras* function, one can search for a certain *sūtras* of the *TK* or certain Buddhist terms within the *sūtras* in real time. In addition, various ways of searching for *sūtras* are provided. One also can easily input search words with the Chinese character input editor. The *sūtras* annotations include a short summary, notes, and philological information for each of the 1,847 *sūtras* comprising the *TK*. One can refer to the annotation of the *sūtras* that one is currently reading. The annotation provides various notes on the *sūtras*, such as the *sūtras* number, Sanskrit title, Pāli title, Tibetan title, author, translator, compiler, abbreviated title, another title, auxiliaries, and references. One also can move backwards or forwards to view the annotation for the previous and following *sūtras* by simply clicking on the arrows instead of searching through every *sūtras*. The dictionary consists of a total of 11,427 terms. The names of the authors and translators as well as titles of the numerous *sūtras* have been digitized and put into the dictionary. Synonyms and various transliterations are also included to

11. A Digital Unified Tripiṭaka project, which compares the *TK* with the *T* and hyper-link the *TK* to the Sanskrit, Tibetan, and Pāli original text, was attempted by Lee Jongcheol (Lee 2000).

enable more practical use of the dictionary. The dictionary also shows how to write the proper nouns of the original language in Korean and even provides the English definitions. Various search options are provided to make it easy for users to look up the terms. The dictionary contains basic information on Chinese characters that can be represented by unicode, which is based on the digitization of the *Dictionary of Variant Chinese Characters*. The *Dictionary of Variant Chinese Characters* contains a total of 29,478 variant characters, corresponding to 7,486 proper characters. The contents of the dictionary include the radical of the character, the total number of strokes, the number of strokes remaining, excluding the radical, the unicode, the Korean, Chinese, and Japanese pronunciations, and the *banjeol*, a method of representing the pronunciation of a Chinese character by two other characters.¹²

In particular, comparison between the *TK* and the *T* will lead scholars to a more precise understanding of East Asian Buddhism than before. The *DTK 2004* will allow more intensified study of the *TK* by providing one-by-one comparisons with the *T*. The one-by-one comparison between the *TK* and the *T* was a solution to the variant characters, and the *Goryeo daejanggyeong gamgyorok* (Comparative Analysis of the TK)¹³ was the result. The *TK* served as the *textus receptus* for the *T* and its modern Chinese counterparts. In particular, the *T* has served among both Eastern and Western Buddhist scholars as the standard collection of Buddhist canonical texts written in Chinese and is today the most widely used document for research on East Asian Buddhism. The distinctive features of the *T* are apparent in its appearance. Each page of each volume of it notes variant readings taken from various editions, ranging in age from the 8th to 18th centuries and in region from China, Turfan, and Korea to Japan. These notes also include corresponding Sanskrit or Pāli words, phrases or sentences when they are available. This series was also published in Western binding so as to facilitate its circulation (Shimoda 2000: 84).

However, scholars have indicated that the *T* contains some inaccurate content. While the *TK* contains over 52 million characters, many of the character forms are ones that are no longer in use today. To resolve this problem, the *T* converted all of the character forms to standard characters in modern usage.

12. These explanations of the *DTK 2004* are based on the information on DISC 1.

13. 49,200 characters among 1,514 *sūtras* of the *TK* were recorded in the *Comparative Analysis of the TK* after comparing them with 1,490 *sūtras* found simultaneously in the *T* by numerous individual checks.

Therefore, it is highly probable that errata exist in the *T* (Hur 2000: 21; Lee 2000: 90; Muller 2000: 39). In the *T*, we can find several classifications of variant characters: the majority of variants in the *TK* were changed to standard forms; A significant number of variants were reproduced exactly as found in the *TK*; and Some of the characters were substituted for new types of variant forms; There are numerous instances in which the incantations expressed with variant forms in the *TK* are expressed with yet a different character in the *T*; and There are many characters in the *T* that have been incorrectly reproduced (Lee 2000: 92-99).

In contrast, the *TK* is the most accurate of all extant *Tripitakas* using Chinese characters (Kim 2002: 172) and owes much to the Monastic Overseer Sugi, who thoroughly studied various texts extant at the time to correct errors and add missing items, thus publishing the *Goryeoguk sinjo daejang gyojeong byeollok* (Record of the Revisions of the *Tripitaka*) from 1246 to 1247. In particular, the *DTK 2004* compares the *TK* with the *T*. The major points of comparison between the *TK* and the *T* in the *DTK 2004* include the differences between the *TK* and the *T*, errors in the *TK*, errors in the *T*, and deformed characters in the *T*. However, a comparison between the two editions, in terms of issues pertinent to the development of East Asian Buddhism, has not yet been conducted. As such, a comparison between the two will contribute to writing new chapters in the field of East Asian Buddhism and, by extension, Asian studies.

3. Potential Contribution to Cultural Sciences

There are many Buddhist texts found only in the *TK*. In particular, twenty works are not covered in the *T* and are found only in the *TK*. Composed from the seventh to the fifteenth centuries, these Buddhist texts will provide important information to better understand the cultural sciences and East Asian Buddhism. These texts include Buddhist lexicons, a Chinese emperor's hymns and poems on Buddhist doctrines, the collation of the *TK* and its catalog, an interpretation of a Zen monk's songs of enlightenment, a genealogical history of Zen Buddhism, an outline of the *Tripitaka*, a collection of public cases of the Zen school, an eminent Korean monk's commentaries on the Flower Garland (Ch. *Huayen*, Kor. *Hwaeom*, Jp. *Kegon*) texts, works on the repentance of transgressions, and the catalog of the supplement to the *TK*. These texts are essential to a better understanding of Buddhist philosophy, bibliographical studies, linguistics and literature, and religious studies (Kim 2002: 160-172), which are still topics for

academic research. For example, *Seonmun yeomsongjip* (Collection of the Meditation School's Enlightened Verses) is preserved only in the *TK*. In-depth research on the *Collection* will contribute to the fields of both East Asian Buddhism and Korean studies, in particular, in terms of bibliography, intellectual history, the relationship between the state and the Buddhist order, and religious dialogue. In addition, a comprehensive investigation of the *Collection* is also expected to clarify thus far unknown aspects of Buddhist history, Seon Buddhism, and literature of Korea (Kim 2000: 69-75).

Challenges to the *DTK 2004*

The current historical information system of Korea has many problems that need to be improved considering the data content and service technology (Kim 2005: 64). With regard to the digitization of sources on the internet general problems are also pointed out. Any text preserved in digital format is as vulnerable as ever-and perhaps even more so. None of data on the internet are guaranteed to survive the next two or three decades. With a different "owner" of the data, often the address changes as well. In most cases, the old links will be "dead." There is also the difficulty to keep texts legally protected or the dangers for the text to be altered or cannibalized without the author's consent (Sobisch 2005). In addition, an increased reliance on digital materials can precipitate a technical divide, putting scholars without requisite computer skills and adequate resources at disadvantage. Search capabilities might also discourage the reading of classical texts and other pertinent sources, thus inhibiting the discovery of new materials and perspectives. That might have a negative impact on research and graduate education, providing disincentive for the careful reading and study of original texts, and perhaps encouraging dilettantism (Poceski 2005). These are also true of the *DTK 2004*.

In particular, the *DTK 2004* has some additional challenges, including the lack of a manual and photo-images, limited access to browsers, no links or mark-ups, and use of mixed Romanization, and no further editorial effort.

1. Lack of a Manual and Photo-images

A user's guide to the *DTK 2004* is available only on DISC 1.¹⁴ However, no printed manual is available at present. Therefore, I was only able to browse the

data in the *DTK 2004* after consulting a program specialist of the RITK. The lack of a printed manual for the *DTK 2004* may make users feel that it is difficult to access.

There are no photo-images of woodblock prints of the *TK* in the *DTK 2004*.¹⁵ However, digitization also serves as a preservation process.¹⁶ Therefore, inclusion of the photo-images of the woodblock prints in their entirety should be a mandatory task for the preservation of the original content of the *TK*.

2. Limited Access to Browsers

The economic stability of Korean Buddhist circles has been primarily maintained by the offerings of lay Buddhist followers, who are mainly lay women in their forties or older. Their donations are primarily aimed at obtaining this-worldly happiness through the Buddha's miraculous power. However, a valuable lesson for the future of Korean Buddhism may be gleaned from the Japanese case, in particular, from Buddhist new religions' approach to electronic media by which the number of Buddhist believers grew (Hoshino 1996: 19-38). As such, the digitization of the *TK* will contribute to increasing the number of Buddhist believers, as well as to encouraging Buddhist studies throughout the world. However, both the general populace and specialists in Buddhism have limited access to the *DTK 2004*.

First, Korean lay Buddhists do not do much reading. In addition, the *DTK 2004* is the digitized version of Buddhist literature in classical Chinese. The general public in contemporary Korea is nearly illiterate in the Chinese language and face difficulty in gaining access to the classics of their own civilization. Therefore, an electronic database of the *TK* translated into the Korean language is essential for easier access. However, this project is now only under consideration.

What is more serious problem is that the *DTK 2004* was published with little consideration for academic use by Buddhist specialists. In fact, although much data have been accumulated, they are not being utilized well by scholars of

14. DISC 1 → *DTK 2004* → Help in the menu bar → Usage F1 → How to Use in Table of Contents.

15. The *DTK 2000* had photo-images of woodblock prints of the *TK*. However, due to their huge volume, the *DTK 2004* did not include these photo-images in its set.

16. Preservation issues have to be taken into account at the very beginning of a digitization project (Hofman 2005: 85).

Korean studies (Kim 2005: 16). This is not exceptional to the *DTK 2004*. According to the people who took the leading part in the digitization of the *TK*, including Daeseok and Chollwan Jeon, their primary purpose was to input basic data with the aim of giving the general public the opportunity to utilize the *TK*. In such a context, the project of publishing the *DTK 2004* was primarily conducted by the RITK members. However, systematic and thorough discourses on this project have seldom been executed in Korean Buddhist academe. As a result, the academic need of Buddhist specialists, the prime end users of the *DTK 2004*, was not fully taken into consideration. It is said that many of digitized databases in contemporary Korea were inputted to meet the purpose of the organization concerned (Gwon 2002). This is also true of the *DTK 2004*. However, the most decisive factor in the digitization of the Korean studies materials is human resources. Cooperative research is the most desired mode for the field to move forward (Kim 2005: 11-5). In addition, the purpose for digitizing data (Eaton 2005: 22) and its most important merit (Choi 2005: 25) is providing enhanced access to those data. Then, limited access to browsers of the *DTK 2004* should be corrected.

3. No Links or Mark-ups

The enhancement of interoperability between databases (Kim 2005: 64) is among the most urgent matters to be addressed. However, the *DTK 2004* has no relevant links, including not having the web release of *The Korean Buddhist Canon: A Descriptive Catalogue* (<http://www.hm.tyg.jp/~acmuller/kor-bud/>), which was completed in November 2004. Published by the University of California Press in 1979, the *Descriptive Catalogue* provides detailed descriptions of the provenance of some 1,600 canonical works, including authors, translators, compilers, dating, geographical sites, monasteries, and equivalent and related titles in Sanskrit, Pāli, and Tibetan in the *TK*. For over two decades, it has been an indispensable source for those requiring information on the texts contained not only in the Korean canon itself, but also in the larger East Asian canon (Muller 2004: 1-3). However, the *DTK 2004* was published without reference to the digitized version of the *Descriptive Catalogue*. In addition, no other relevant links, such as the *Hanguk Bulgyo jeonseo* (Collected Works of Korean Buddhism) (<http://ebti.dongguk.ac.kr/ebti/main.html>), Resources for East Asian Language and Thought (<http://www.acmuller.net>), or the Center for Buddhist Studies at UCLA (<http://www.isop.ucla.edu/buddhist>), which are closely related

to studies on Korean Buddhism and, by extension, East Asian Buddhism, are not provided in the *DTK 2004*.

In particular, published by Dongguk University Press in thirteen volumes of photolithographic prints from 1979 to 2001, the *Collected Works of Korean Buddhism* contains extant Buddhist works compiled or written by Korean scholar-monks from the seventh to the twentieth centuries. They are an excellent source for the study of Korean Buddhism and show the way Koreans have adopted Buddhism. The first fascicle of the collection begins with works by Wonhyo (617-686), the most creative thinker in the history of Korean Buddhism. The electronic input of his entire *oeuvre*, which consists of about 700,000 characters, was completed in China in 1993. As of 2005, twelve volumes of the *Collected Works of Korean Buddhism* were digitized. Therefore, the link of the *DTK 2004* to the digitized *Collected Works of Korean Buddhism* will greatly contribute to identifying the major intellectual concerns of Korean Buddhist scholars throughout Korean history.

In addition, a database project should have mark-ups (Yi 2000: 172) and the vitally important task of mark-ups must be carried out thoroughly and accurately (Muller 2004: 2). Based on the digitized version of the *Ershi wu shi* (Twenty Five Historical Works) of China, the digital *Veritable Records of the Joseon Dynasty* already had mark-ups (Kim 1995: 71-83; Yi 2000: 174-5). However, the *DTK 2004* currently has none.

4. Use of Mixed Romanization

There are largely two Romanization systems for the Korean language. One is the McCune-Reischauer system and the other is the revised Romanization system. The former was a product of the 1930s and has been most widely used by overseas scholarship. The latter was newly devised by the Ministry of Culture and Tourism of Korea in 2000 and is being used in Korean academe. One of these two systems should be chosen and applied consistently, without mixing the two or landing somewhere in between. However, the Romanization system applied to the *DTK 2004* is not consistent. For example, “Changgyong Panjon” in the *DTK 2004* should be “Changgyŏng p’anjŏn” according to the McCune-Reischauer system or “Janggyeong panjeon” in accordance with the revised Romanization system.

5. No Further Editorial Effort

No editorial effort was visible in the digitization process of Tibetan sources (Sobisch 2005). This is also true of the *DTK 2004*. It is the final version of its kind at present and people concerned testified that no more funding was available for further technical advancement or editorial work. This is a serious problem for the ongoing and appropriate use of the digital version by end users. First of all, hardware and software products are being upgraded and replaced roughly every eighteen months. Companies in the information technology sector have reported that the majority of products and services they offer did not exist five years ago. Therefore, for cost-effectiveness, hardware and software products have to be changed from generation to generation (Chen 2005: 54). In addition, when digitization first became popular it was assumed that it could serve as a preservation process. However, it is said that digitization must not be viewed as being a substitute for other preservation activities. Managers of digital collections need to ensure that their digital files remain accessible over time. The ongoing maintenance means data entry and data cleaning, ensuring the usability of access applications, providing support for end-users, upgrading server hardware and operating system software over time, maintaining server security and ensuring that restoration of applications and data from backups is possible. This requires an organizational commitment to continued maintenance, which will result in ongoing costs (Eaton 2005: 21). Otherwise, the digital storage of texts will be in no way more secure over time than would simply printing it even on acid-free paper and, indeed, a printed text may have to be given precedence over a digitized version (Sobisch 2005).

Future Directions for the *DTK 2004*

Will the digitization of the *TK* have any bearing on the possibilities of its returning to a level of prominence and usage equal to that of the *T*? In short time, such a question will become moot. As far as authoritativeness or accuracy are concerned, there is no doubt about the status of the Korean canon, and therefore, if it is made readily available in an easy-to-use format, consulting it could well become standard practice. However, whether or not one will want to consult the digitized *TK* depends primarily on accessibility and usability (Muller 2000: 40-1). Therefore, for better accessibility and usability of the *DTK 2004*, the follow-

ing issues should be further taken into consideration: Participation of a wide range of specialists; establishment of better planning; setting up of standardization, maintenance of high quality; and applying economics to the *DTK 2004*.

1. Participation of a Wide Range of Specialists

It was already pointed out that the digitization of historical sources on Korea should be carried out by the cooperation of experts in both history and computer science (Yi 2000: 163). Likewise, the future revision and maintenance of the *DTK 2004* should be conducted by the cooperation of a wide range of sincere experts in both Buddhist studies and computer science. In addition, patience and perseverance are also important factors for the successful result of the *DTK 2004* as demonstrated as follows:

We would emphasize that only those scholars blessed with unusual qualities of patience and perseverance, and willing to wait a long time for results and recognition, should attempt such a project [the Munkwa project]. (Wagner 1982: 144)

2. Establishment of Better Planning

Any digitization project needs careful thinking and planning before initiation. This is because the available digital data, when used prematurely, may cause more damage than benefit (Sobisch 2005). “The Munkwa project” took three years in preparation for the digitization of data and fifteen years more for actual digitization. Nevertheless, its results have not come out yet. The reason was because the scholars concerned wanted to provide more exact and better information. In contrast, the digitization projects of historical sources of Korea have tended to focus on quantity rather than on quality (Choe 2001: 401). By extension, the digitization projects carried out by the academic circles of Korea have been primarily interested in gaining hasty results rather than promoting research activities (Park 2004: 262). The digitization of Tibetan data also emphasized quantity rather than quality, which frequently led to the hurried publication of electronic resources. The result is sometimes disastrous and often discouraging from a scholarly point of view (Sobisch 2005). The *DTK 2004* is also, in part, subject to such criticism.

In particular, from the very start consideration has to include user needs

(Hofman 2005: 83-84).¹⁷ The digitization projects of historical sources aimed to meet the needs of both historians and the general public (Choe 2001: 399). However, the *DTK 2004* primarily focused on inputting basic data. Instead, little consideration was given to its major end users, i.e., academics, including Buddhist specialists.

3. Setting up Standardization

Existing digital repositories lack standardization (Park 2005: 106). However, the most important issue related to digitization and database construction on Korean records heritage is to set up the standardization of digitization processes and facilities. It is urgently necessary to develop comprehensive standard systems for digitization (Kim 2005: 258). This is also true of the digitization of the *TK*, and the future revision of the *DTK 2004* needs to take this issue into consideration. In addition, the cataloging of a given subject, including its researchers, and the treatment of characters used only in specific works also need further consideration.

4. Maintenance of High Quality

In the decision making process, managers remember that the digitized images should be high quality, enough to meet an institution's needs and user requests (Park 2005: 109). However, it was pointed out that many conventions of careful and critical scholarship in philology and lexicology have been dispensed with by producers of Tibetan electronic texts and dictionaries (Sobisch 2005). In addition, electronic data needs to be occasionally transferred due to IT advances and media changes. In any circumstance, electronic data should guarantee the authenticity, integrity, originality of data, consistent format, process, and data reliability (Yoon 2005: 133). These indications also, in large part, apply to the *DTK 2004* and remain for further consideration.

5. Economics of the *DTK 2004*

The majority of challenges to the *DTK 2004* are concerned with cost. Therefore,

17. "The Technical Advisory Service for Images" and the "NINCH Guide to Good Practice in Digital Representation and Management" are useful guidance for setting up a digitization project (Hofman 2005: 84).

the official securing of an ample budget is of prime importance for its prospective future. Academic databased serve as an important infrastructure for the progress of scholarship in Korea and it should be carried out in terms of public works (Kim 1995: 91) under the continuous sponsorship of the government (Choe 2001: 402). We may learn a lesson from the Electronic Records Archives program, the strategic response of the United States National Archives and Records Administration to the challenges posed by electronic records. This program is dependent upon continued congressional funding and both the President and Congress have looked favorably on this project for the past seven years since 1998 (Eaton 2005: 23-27). There is a saying that the equality of records management in a country reflects the advancement of democracy in that country. The Korean government also set records management as a primary goal of Government Innovation Tasks and has led the reform movement since the mid-1990s (Park 2005: 36-38). As a result, the information project sponsored by the Korean government from 1998 to 1999 accelerated the digitization of historical sources of Korea (Choe 2001: 388; Kim 2005: 45). The *TK* is a National Treasure of Korea and the Depositories of the *TK* (*Janggyeonggak*) are now registered on the UNESCO world heritage list. Therefore, the Korean government must take measures to fund digital preservation, organizational process, technology continuity, and the upgrading of the *DTK 2004*.

Conclusion

This paper examined the history of the digitization of sources for Korean studies, the content of the *DTK 2004*, benefits from it, challenges to it, and future directions for it in the study of East Asian Buddhism. The *DTK 2004* hopes contribute to promoting Buddhist studies throughout the world. However, it has both benefits and challenges. Its strengths include the revival of the original text, more useful functions such as comparison between the *TK* and the *T*, and its potential contribution to the cultural studies of East Asia. It also has some challenges, including the lack of a printed manual or photo-images of the woodblock prints of the *TK*, limited access to browsers, no links or mark-ups, the use of mixed romanization, and no further editorial effort. However, if we want better digital resources and the more responsible use of them, we must first resolve these problems. Furthermore, issues of participation of a wide range of specialists, the establishment of better planning, the setting up of standardization, the

maintenance of high quality, and the economics of the *DTK 2004* remain to be resolved. In particular, the majority of challenges to the further usability of the *DTK 2004* are primarily concerned with people in charge and cost. Buddhist specialists, computer scientists, and other related scholars should cooperate for the lasting editorial works of the *DTK 2004*. In addition, fundraising for the constant maintenance and technical advancement of the *DTK 2004* is the most urgent impending issue. Government funding is of prime importance for the ongoing preservation and accessibility of the *DTK 2004*.

References

- Arnold, Mattias. 2005. "Historical Inscriptions at Cloud Dwelling Monastery: Preparations for a CD-ROM." Paper presented in the session, "Windfalls and Pitfalls; The Use of Digitized Source Material in the Study of Asian Buddhism," organized by Morten Schlutter, at the AAS Annual Meeting, held at Hyatt Regency Chicago, Chicago, U.S.A., on March 31-April 3.
- Buswell, Robert E., Jr. 1989. "Haein-sa." *Korean Culture* 10(1), *the First Twenty Years: 1980~1999 Issues* [CD-ROM version].
- Chen Su-Shing. 2005. "Practical Technology for Digital Archives and Preservation." In *Digital Preservation of Cultural Heritage Resources: Implementation, Standardization, and Strategies*, The 1st International Symposium in Celebration of UNESCO/Jikji Memory of the World Prize, hosted by the City of Cheongju, Korea, on September 1-2, pp. 54-63.
- Choi Jin Ok. 1994. "Hanguksa jaryoui jeonsan cheori" (Digitization of Sources on Korean History). *Jeongsin munhwa yeongu* (Korean Studies Quarterly) 17(3): 71-89.
- . 2001. "Hanguk yeoksa jaryoui jeonsanhwa hyeonhwanggwa geu banghyang" (The Present and Direction of the Digitization of Historical Materials on Korea). Paper presented at the International Conference on National Literature of Korea entitled *Isibil segi Joseon minjok munheon ui balgul gwa yeongu* (The Excavation and Study of National Literature of Korea in the Twenty First Century), held in Yanji, China, on September 21-22, pp. 385-402.
- . 2005. "The Information of Korean Historical Data." *The Review of Korean Studies* 8(4): 19-42.
- Eaton, Fynnette. 2005. "Digitization and Preservation in the United States." In

Digital Preservation of Cultural Heritage Resources: Implementation, Standardization, and Strategies, The 1st International Symposium in Celebration of UNESCO/Jikji Memory of the World Prize, hosted by the City of Cheongju, Korea, on September 1-2, pp. 21-8.

- Goryeo Daejanggyeong Yeonguso (Research Institute of Tripitaka Koreana), ed. 2004. *Goryeo daejanggyeong gamgyorok* (The Comparative Analysis of the Tripitaka Koreana). Seoul: Goryeo Daejanggyeong Yeonguso (<http://www.sūtra.re.kr>).
- Gungnaewae hangukhak jaryoui bojeon siltaewa jeonmang* (The Actual Conditions of Preservation of Sources for Korean Studies at Home and Abroad). Recited from Gwon Hyeongjin, “Gungnaewae hangukhak jaryodeul eotteoke jeongni iyong hana” (How to Arrange and Use Materials on Korean Studies at Home and Abroad). *Budapia* (<http://www.buddhapia.com>) (December 11, 2002).
- Gwon Nae-hyeon. 2004. “Materials on Korean Studies: Digitalization of and New Research Trends in Hojeokdaejang from the Late Joseon Dynasty.” *The Review of Korean Studies*, 7(4): 229-46.
- Ha Jongun. 2005. “Jalmot sseuin ‘sokjanggyeong’ myeongching sae hakkie bakkwinda” (The Mis-Title of the *Sokjanggyeong* Will Change in the New Term), *Bulgyo sinmun* (Buddhist Newspaper), January 11.
- Hofman, Hans. 2005. “Improving the Availability of Cultural Heritage Sources: Developments in Europe.” In *Digital Preservation of Cultural Heritage Resources: Implementation, Standardization, and Strategies*, The 1st International Symposium in Celebration of UNESCO/Jikji Memory of the World Prize, hosted by the City of Cheongju, Korea, on September 1-2, pp. 80-91.
- Hur In-Sub. 2000. “The Academic Significance of and Future for the *Digital Tripitaka Koreana* 2000.” Paper presented at the International Conference on Buddhism and the 21st Century Digital Information Society, sponsored by The Research Institute of Tripitaka Koreana, held at Dongguk University, Seoul, Korea, on December 7-9, pp. 21-9.
- Jang Yeongseop. 2002. “*Hanguk bulgyo jeonseo* CDro bonda” (The Digital Collected Works of Korean Buddhism). *Bulgyo sinmun* (Buddhist Newspaper), July 24.
- Kim Byongsun. 2005. “Guest Editor’s Introduction Digitalization of Korean Studies Materials: An Overview.” *The Review of Korean Studies* 8(4): 5-17.
- . 1994. “Hangukhak jaryo jeonsan cheoriui gibon munje” (Basic Issues

- in the Digitization of Sources on Korean Studies). *Jeongsin munhwa yeongu* (Korean Studies Quarterly) 17(3): 3-33.
- Kim Hyeon. 1988. "Dongyanghak jaryo jeonsanhwaui gusanggwa gwaje" (The Scheme and Theme of Computerized Data on East Asian Studies). *Minjok munhwa yeongu* (A Study of Korean Culture) 21: 291-323.
- . 1995. "Hanguk gojeonjeogui jeonsanhwaui seonggwawa gwaje— Joseon wangjo sillok CD-ROM gaebal saebui gyeonggwawa baljeon banghyang" (Results and Problems of the Digitization of Korean Classics— Progress and prospects of the CD-ROM project on the *Veritable Records of the Joseon Dynasty*). *Minjok munhwa* (National Culture) 18:70-102.
- . 2005. "Historical Information Systems Administered by the Korean Public Sector -Present and New Challenges-", *The Review of Korean Studies* 8(4): 43-69.
- Kim Jongmyung. 2000. "The Digitization of the *Tripitaka Koreana* and the Collection of the Meditation School's Enlightened Verses." Paper presented at the International Conference on Buddhism and the 21st Century Digital Information Society, sponsored by The Research Institute of Tripitaka Koreana, held at Dongguk University, Seoul, Korea, on December 7-9, pp. 69-73.
- . 2002. "The *Tripitaka Koreana*: Computerization and Significance for the Cultural Sciences in a Modern Globalized World." In *Korea and Globalization: Politics, Economics, and Culture*, edited by James Lewis and Amadu Sesay. London: RoutledgeCurzon, pp. 154-81.
- . 2005. "The Digital *Tripitaka Koreana* 2004." Paper presented in the session, "Windfalls and Pitfalls: The Use of Digitized Source Material in the Study of Asian Buddhism," organized by Morten Schlutter at the AAS Annual Meeting, held at Hyatt Regency Chicago, Chicago, U.S.A., on March 31-April 3.
- Kim Sung-Soo. 2005. "Preservation of World Records Heritage in Korea and Further Registry." In *Digital Preservation of Cultural Heritage Resources: Implementation, Standardization, and Strategies*, The 1st International Symposium in Celebration of UNESCO/Jikji Memory of the World Prize, hosted by City of Cheongju, Korea, on September 1-2, pp. 257-60.
- Lancaster, Lewis R., compil. 1979. *The Korean Buddhist Canon: A Descriptive Catalogue*. Berkeley: University of California Press.
- Lee Jongchol. 2000. "Case Studies in Constructing a Unified *Tripitaka*." Paper presented at the International Conference on Buddhism and the 21st Century

Digital Information Society, sponsored by The Research Institute of Tripitaka Koreana, held at Dongguk University, Seoul, Korea, on December 7-9.

- Lee Kyoo-Kap. 2000. "The Relationship between the Chinese Character Forms in the Goryeo and the Taisho *Tripitakas*: With emphasis on the reading of variant forms." Paper presented at the International Conference on Buddhism and the 21st Century Digital Information Society, sponsored by The Research Institute of Tripitaka Koreana, held at Dongguk University, Seoul, Korea, on December 7-9, pp. 90-9.
- Lee Kyoo-Kap, ed. 2000. *Goryeo daejanggyong ichejajeon* (Dictionary of Variant Chinese Characters in Tripitaka Koreana). Seoul: Goryeo Daejanggyeong yeonguso.
- Muller, Charles. 2000. "Buddhist Canonicities: Value and Evaluation in the Digital Age on the Occasion of the Celebration of the Publication of the Digital *Tripitakas Koreana*." Paper presented at the International Conference on Buddhism and the 21st Century Digital Information Society, sponsored by The Research Institute of Tripitaka Koreana, held at Dongguk University, Seoul, Korea, on December 7-9, pp. 37-43.
- . 2004. "Digital Editor's Preface to the Web Publication of the *Descriptive Catalogue*," <http://www.hm.tyg.jp/~acmuller/>, October 20, pp. 1-18.
- Park Chan-Woo. 2005. "Future Directions and Vision for Government Records Management Innovation." In *Digital Preservation of Cultural Heritage Resources: Implementation, Standardization, and Strategies*, The 1st International Symposium in Celebration of UNESCO/Jikji Memory of the World Prize, hosted by the City of Cheongju, Korea, on September 1-2, pp. 36-42.
- Park Eun G. 2005. "Imaging and Metadata Standards in Digitization: Practical Strategies." In *Digital Preservation of Cultural Heritage Resources: Implementation, Standardization, and Strategies*, The 1st International Symposium in Celebration of UNESCO/Jikji Memory of the World Prize, hosted by City of Cheongju, Korea, on September 1-2, pp. 105-15.
- Park Soohwan. 2004. "Gurye ryussiga sanghwal ilgi Kieo-HRAFreul iyonghan saenghwalsa yeonguui gicho yeongu jaryo hwakboreul mokjeogeuro" (A Study on the Method of Computerization of Gurye Yu's Diary-for the Purpose of Setting Basic Sources of Life History with HRAF). *Gomunseo yeongu* (A Study of Ancient Documents) 25 (8): 259-89.
- Poceski, Mario. 2005. "Use of Digitized Sources in the Study of Chan." Paper

- presented in the session, “Windfalls and Pitfalls; The Use of Digitized Source Material in the Study of Asian Buddhism,” organized by Morten Schlutter at the AAS Annual Meeting, held at Hyatt Regency Chicago, Chicago, U.S.A., on March 31-April 3.
- Schlutter, Morten Schlutter. 2005. “Windfalls and Pitfalls: The Use of Digitized Source Material in the Study of Asian Buddhism.” Abstract for Proposed Panel at the AAS Annual Meeting, March 31-April 3.
- Shimoda, Masahiro. 2000. “The Significance of the Publication of the *Taisho Tripitaka* in Modern Buddhist Studies.” Paper presented at the International Conference on Buddhism and the 21st Century Digital Information Society, sponsored by The Research Institute of Tripitaka Koreana (December 7-9), pp. 84-6.
- Sobisch, Jan-Ulrich. 2005. “Digitization in Tibetan Studies: Problems and Remedies.” Paper presented in the session, “Windfalls and Pitfalls; The Use of Digitized Source Material in the Study of Asian Buddhism,” organized by Morten Schlutter at the AAS Annual Meeting, held at Hyatt Regency Chicago, Chicago, U.S.A., on March 31-April 3.
- Wagner, Edward W. 1982. “Problems in the Computerization of Materials in the Korean Studies Field: A Report on the Munkwa Project.” *Bogo nonchong 82-3 Hangukhak jaryoui jeonsanhwa yeongu* (Report 82-3 A Study of the Digitization of Materials on Korean Studies). Seongnam: The Academy of Korean Studies, pp. 127-44.
- Yi Namhui. 2000. “Joseon sidae jaryoui jeonsanhwa-database guchugui hyeondangyewa gwaje” (Digitization of Materials of the Joseon Period—The Present of and Prospect for Their Database). *Joseon sidae sahakbo* (Journal of History of the Joseon Period) 12 (3): 153-80.

Kim Jongmyung received his doctoral degree in Buddhist studies from the University of California, Los Angeles, and is currently Assistant Professor of the Program of International Korean Studies at The Academy of Korean Studies. He is interested in Zen thought and applied Buddhism.