

The Asian Classics Input Project (ACIP): Past, Present and Future

Robert Chilton
ACIP, USA

Introduction

Started in 1987, the Asian Classics Input Project (ACIP) is an ongoing effort to preserve and disseminate important classical Asian literature in digital format. Approximately 45,000 pages have been input from woodblock prints of the Kangyur and Tengyur collections of classical Sanskrit literature in Tibetan translation. The works in these collections, dating from the period of 500 BCE up to 900 CE, set forth some of the most significant ideas of Asian thought and culture. ACIP has also input more than 75,000 pages of native Tibetan writings, based on these two collections and dating from 1000 CE to the present, which cover such topics as philosophy, ethics, logic, epistemology, psychology, hermeneutics and metaphysics, among others. The Project has cooperated with local institutions to create comprehensive electronic catalogs of the extensive but relatively inaccessible Tibetan collections located in St. Petersburg, Russia and Ulaanbaatar, Mongolia. Over 70,000 separate titles have been cataloged to date, with work projected to continue for another decade or more. ACIP materials are distributed without charge on diskette, CD-ROM, and via the Internet.

ACIP faces a number of challenges at present. Because so few of the books reached India following the Tibetan diaspora of the 1950s and 1960s, we now have to look for important texts in other places, such as in the libraries of Russia, Mongolia, and the United States. Continuing input of the Kangyur collection will depend on obtaining a copy of an appropriate edition. Finally, although key operations personnel continue to serve the Project, we are temporarily without our founder and principal fundraiser, Geshe Michael Roach, who is in strict retreat until 2003.

For the future, we can expect that the revolution in scholarship brought about by the availability of primary source materials in digital format will continue as people find new ways to take advantage of the efficiencies of preservation, access, and searchability engendered by digitalization. The explosion of Tibetan-language books published by educational institutions in South Asia using computerized typesetting of electronic texts will likely continue and spread. Specific ACIP plans include: continue and expand text input and cataloging efforts; integrate catalog information for each text directly into the text itself using XML markup; transition from

ASCII (roman transliteration) to Unicode; enhance the accessibility and usability of ACIP materials; and cooperate with other organizations to provide a broad and coherent selection of digital resources to end-users.

Technologies and especially international standards have evolved in recent years, enabling new modes of accessing and sharing digital resources. At the same time, the sheer quantity of digitized material calls for an assessment as to which formats are most appropriate for storing and disseminating ACIP data. Below I will present an overview of the various ACIP digital resources, with examples of each. But first I will briefly mention some specific trends and developments that are impacting the creation and use of digital data in general and Tibetan language materials in particular.

Data encoding issues

Due to the absence of a universal encoding standard for Tibetan, display of data in Tibetan script has necessitated using one or another of the various proprietary Tibetan fonts. Since all these different fonts have their own way of encoding the Tibetan characters, a user must have access to the particular Tibetan font(s) used in the document in order to view it. Moreover, if any changes are made to the mapping of the font, the data is at risk of being corrupted.¹ Because ACIP is committed to maintaining its data in a non-proprietary format useable on any computer, all Tibetan language materials have been transliterated into standard Latin (ASCII) characters. However, in the near future we can expect widespread adoption of Unicode (ISO-10646)—which will enable us to represent Tibetan-script data directly in a universally readable format.²

Document format issues

Beyond Unicode, a second significant development is seen in the area of standards that define the format of documents. At present there are three basic categories of file format for Tibetan. A “plain-text” format is generally used with materials encoded in roman transliteration.

¹ This is not merely a theoretical issue; mappings for Tibetan fonts have been changed in the past and these changes, often associated with changes in the overall format of the data (i.e., in a proprietary word processing system), have required tedious conversion of existing Tibetan data. Since ACIP Tibetan data is transliterated, it has escaped this particular problem.

² Tibetan first appeared in Unicode version 2.0 (1996) and was expanded in version 3.0 (2000). Efforts are now underway to support Unicode Tibetan on various platforms; the *Tibetan Language Kit* for Macintosh already supports Unicode (albeit in a non-normative manner).

Materials encoded using Tibetan-script fonts are composed in either a proprietary word processor format or in a more generic format, such as RTF or HTML. Although HTML is merely one instance of SGML, as the driving force behind the World Wide Web it has revolutionized the Internet and the manner in which we now access and share information. However, because HTML is oriented toward formatting for display rather than markup of content, we have not migrated ACIP data to HTML. Rather, ACIP has taken a lowest-common-denominator approach, storing and disseminating all data in a “plain-text” format.

The promise of XML

In a single stroke, XML promises to remove both technical limitations (of encoding and format) that have plagued the field of Tibetan studies. Although EBTI hosted experts on TEI back in 1994, only the advent of XML (which incorporates Unicode) makes implementation of such markup truly practical for Tibetan-script materials. Migrating existing Tibetan digital resources to Unicode and XML will be more or less difficult depending on the format of the source data. Conversion of data that employs roman transliteration (such as ACIP format) will not be difficult; in fact, we already have a Java applet that converts ACIP files to Unicode. In contrast, a font-encoded document, particularly one that resides in a proprietary word processor format, will be a challenge to convert.

ACIP Digital Resources: Overview

The Project’s major work is the creation of **digital texts**. Each book is input twice in roman transliteration by independent typists and an automated computer program is then used to compare the two typed versions, letter by letter. This work is accomplished in over a dozen locations by Tibetan refugees who receive training, equipment, and modest wages from the Project. Since 1987, more than 2,000 separate titles comprising some 120,000 pages of accurate and searchable text have been input, archived, and freely distributed to thousands of users worldwide.

Each of the 2,000-plus works in the text database is assigned an ACIP catalog number and entered into the **ACIP Master Catalog**. These entries record the title of the text (in multiple languages if appropriate), the author, the edition, volume and folio pages, the size of digital file size in bytes, and other relevant information.

In order to locate and obtain books for input, ACIP searches through the surviving collections of woodblock printed texts and manuscripts in places such as Russia and Mongolia where we create **digital catalogs** of the materials. Cataloging work began over seven years ago in Russia (where more than 70,000 titles have been cataloged) and has recently begun in Mongolia (where 3,500 titles have been cataloged).

For the benefit of scholars and researchers of this material, ACIP has also created digital versions of native catalogs, dictionaries, lexicons, and other important **reference materials**. Please refer to the following section for example entries from the *Great Tibetan Dictionary* and from a Sanskrit-Tibetan dictionary—two very useful reference works input by the Project.

1. Digital texts

The digital texts are marked up in two different ways. First, the texts themselves embed information that explicitly identifies the title of the work in Tibetan (and Sanskrit, if applicable), the author (and translators, if applicable) and the various sections (BAM PO). Second, ACIP provides rudimentary markup for indicating pagination and line-breaks in the original edition, change of type size (YIG CHUNG), special syllable-marks, non-Tibetan material in bilingual documents, comments and corrections, and the presence of ornamental script or images (with or without captions).

2. ACIP Master Catalog

The ACIP Master Catalog plays a vital role in organizing and providing access to the various digital resources created by the Project. For example, the Master Catalog relates the ACIP catalog number for a resource with the database file name for the most current version of that resource. With regard to textual materials, the Master Catalog records essential information about title, author, volume, edition, verification level, and so forth.

The basis of ACIP catalog number for text files is the Tohoku University catalogs of the Kangyur, Tengyur, and native Tibetan works. Whenever a work has already been assigned a number in one of these catalogs, this becomes the basis of the file name. When a work has not been assigned such a number (normally for a text in the Sungbum Collection), it receives the next available ACIP number. The initial letter, K, T, S, or R, serves to identify in which general collection a text resides (i.e., **K**angyur, **T**engyur, **S**ungbum, or **R**eference).

No attempt is made to assure that even successive titles in the collected works of a single author receive successive catalog numbers, since this becomes a moot point once we are released from the limitations of a printed paper catalog. All the catalog entries are loaded into the catalog database, and can be sorted on author with a single click of the mouse button. The “Volume/section” field in the database then tells you where the work was originally located within a particular volume, even if the volume was not sequentially paginated.

It is our intention for the future to use XML to integrate most or all of the Master Catalog information for a given text into the text document itself. (As previously mentioned, this step is contingent upon support for Tibetan-script materials within Unicode computing environments.) But even with the texts marked up in this manner, a separate Master Catalog will still be useful and perhaps even indispensable. It should be possible, however, to automatically derive the desired information directly from the marked up texts.

3. Cataloging Projects

In their present form, each digital catalog created by the Project is a simple flat-file database: straight text in the universal ASCII code that can be read and searched on any computer. The raw data of the catalogs are expressed mainly in single-letter codes. This system allows the input operators to catalog more items per day and also has the advantage of allowing the entire catalog to be easily loaded into a relational database, which enables sorting on any field.

Each title is treated as a separate work in its own right, and complete cataloging information is given for each one. Since the catalog is computerized, it is possible to include a level of detail not achieved in any previous printed catalogs. Therefore, a total of twenty-two different details of each title are recorded, including the entire colophon—containing composition and printing information along with valuable information on the Tibetan teaching lineages and other historical information.

The cataloging projects in St. Petersburg and Ulaanbaatar have generated an intriguing side-effort, which is a graphics collection of all the exquisite monastic and other seals found stamped on the woodblock manuscripts of the collections. The presence of a seal is recorded in each catalog, and the seals each have their own unique identification number under the "GSP" and "GUB" file series in the ACIP Graphics Collections. The seals, as well as carved illustrations of eminent teachers throughout history located in the collections and in other texts input by ACIP, are indexed and cataloged in a format that allows them to be integrated into the ACIP

Master Catalog.

3a. Cataloging in St. Petersburg

Two of the greatest collections of Tibetan woodblock prints exist far from Tibet, in St. Petersburg, Russia. The first collection, at the St. Petersburg Branch of the Institute of Oriental Studies of the Russian Academy of Sciences, contains approximately 25,000 volumes, representing up to 200,000 titles—although we do expect a good number of repeat titles by the time we are finished. The second collection is found at the Oriental Library, a part of the Library of the University of St. Petersburg. It is a carefully selected treasure of 3,300 separate titles. The books in each of these collections have been faithfully collected by generations of Russian explorers and scholars ever since the time of the tsars, and maintained with devotion by library staff even in the most difficult of times, such as the three-year siege of St. Petersburg during World War II.

The catalog as it stands now contains over 70,000 entries; it is anticipated that the cataloging work alone (not to mention input or imaging of the texts) will continue for a decade or more. ACIP has made a commitment during this time to assist the Tibetan Collection of the Institute of Oriental Studies in setting up reproduction facilities that will allow the Institute to fulfill orders from interested persons for copies of texts they need to work with. It is hoped that the existence of this reproduction facility, and the fees paid to the Institute by scholars for copies, will help defray the considerable cost of maintaining this precious treasure for future generations.

3b. Cataloging in Ulaanbaatar

We are pleased to report that cataloging of the extraordinary Tibetan-text collections of Mongolia has commenced under similar arrangements as those made with the Institute in St. Petersburg. We estimate that there are approximately 1.4 million titles in the two collections housed in Ulaanbaatar, although a good number of these titles are, no doubt, duplicates. The Mongolian collections are especially important for the Project, since we have found many texts here that we have not found anywhere else in the world.

The library of the main Buddhist monastery of the country, Gangdan Tekchen Ling, holds a vast collection of books that, during the religious persecution of the communist era—from 1937 until only recently—were not properly cared for. We estimate that it will take many decades

simply to put all the loose woodblock pages of the works back into order and catalog them properly. The efforts by local Mongolian monk scholars to restore the library are admirable, although they have a serious lack of material and financial support needed to complete their task.

The collection at the Mongolian National Library is in much better condition; it generally contains woodblock prints and manuscripts that the former communist authorities determined had monetary value, and was better cared for. Although some of the materials have already been cataloged, these entries are written on paper and generally catalog only entire volumes rather than the individual works in each volume. The library is enthusiastically cooperating with ACIP staff in creating a comprehensive electronic catalog wherein each separate work is given as its own entry, complete with colophon information, as in the catalogs of the Tibetan collections in St. Petersburg and at Gangdan Tekchen Ling.

4. Native catalogs, dictionaries, lexicons, and other reference materials

The Project has input (or otherwise obtained) a large variety of digital reference materials including the complete Tibetan and Sanskrit catalog listings of the U.S. Library of Congress, various collections of graphics images, native catalogs to multiple editions of the Kangyur and Tengyur, the Tohoku University catalog of native Tibetan works, a comparative edition of the *Mahavyupatti*, Vostrikov's *Tibetan Historical Literature*, Whitney's *Sanskrit Grammar* and his *Sanskrit Verbal Roots*, and many other important works such as the two dictionaries excerpted in the following section.

ACIP Digital Resources: Examples

Example 1: Text excerpt from ACIP catalog number TD3809 (file name TD3809M.ACT)

[translation of the following portion in bold typeface:] In the language of India: *Ārya Prajñā Pāramitā Saṅgraha Kārikā*; in the language of Tibet: *A Brief Presentation of the Exalted Perfection of Wisdom, Set in Verse*.

#, ,,RGYA GAR SKAD DU, A'ARYA PRADZNY'A P'A RA MI T'A SAm GRA HA

K'A RI K'A, BOD SKAD DU, 'PHAGS PA SHES RAB KYI PHA ROL TU PHYIN MA

BSDUS PA'I TSIG LE'UR BYAS PA, BCOM LDAN 'DAS MA SHES RAB KYI PHA
ROL TU PHYIN MA LA PHYAG 'TSAL LO, ,SHES RAB KYI PHA ROL TU PHYIN
GNYIS MED, ,

YE SHES DE NI DE BZHIN GSHEGS, ,BSGRUB BYA DON DE DANG LDAN PAS,
,GZHUNG LAM DAG LA DE'I SGRA YIN, ,RTEN DANG DBANG DU BYA BA
DANG, ,LAS NI SGOM PA DANG BCAS DANG, ,RAB DBYE RTAGS DANG LTUNG
BA DANG, ,PHAN YON BCAS PAR YANG DAG BRJOD, ,DAD LDAN 'JUG PA'I
YAN LAG

TU; ,STON PA DBANG PO 'KHOR DANG NI, ,YUL DUS DAG KYANG NGES
BSTAN PA, ,SDUD PO RANG NYID TSAD MAR NI, ,RAB GRUB PHYIR YIN 'JIG
RTEN NA, ,YUL DANG DUS KYIS NYER MTSON ZHING, ,DBANG POR BCAS PA
SMRA BA YIN, ,SMRA PO TSAD MAR RJES SU RTOGS, ,TSIG PHRAD BDAG
THOS

LA SOGS PA, ,DI DAG THAMS CAD ZHAR LA NI, ,BRJOD PA YIN TE GTZO BO'I
DON, ,SUM CU RTZA GNYIS DE DAG NYID, ,STONG PA NYID KYI DBYE BA NI,
,RNAM PA BCU DRUG BRGYAD STONG PAR, ,RIM PA JI BZHIN BRJOD PA NI,
,GZHAN GYIS BSTAN PAR SHES PAR BYA, ,DE LTAR JI SKAD BSHAD

@293A *, ,DON GYIS, ,BRGYAD STONG 'DIR DON MA TSANG MED, ,DIR
GZHUNG BSDUS PA YIN 'DOD DE, ,DON NI JI SKAD BRJOD NYID DO, ,BYANG
CHUB SEMS DPA' DE NYID DU, ,NGAS MA MTHONG ZHES THUB PAS GSUNGS,
,ZA PO NANG GI DNGOS RNAMS KYI, ,STONG PA NYID NI DE YIS BRJOD, ,

.....

@294B DE YI KHYAD PAR RNAM RTOG BSAL, ,GANG PHYIR MING BCOS
BRJOD BYA YI, ,CHOS DE DAG KYANG BRTAGS YIN PAS, ,SGRA DON DAG GI
'BREL PA NI, ,NGO BO NYID DU DES MI BZHED, ,BYIS PA'I PHYI ROL DON DU
NI, ,MNGON ZHEN 'KHRUL PAS SPYAD PA YIN, ,THA SNYAD 'DI YANG DE
BZHIN TE, ,DI YANG DE BZHIN TE, ,DI LA DON

'GA' YOD MA YIN, ,DES 'DIR JI LTAR MING BYA BA, ,DE LTAR BRJOD BYA'I
DNGOS PO DE, ,YOD MIN DANG PHYIR MING JI BZHIN, ,DON DU RTOG PA
'DOD MA YIN, ,BDEN DON RNAM RTOG SEL BA NA, ,SHES RAB PHA ROL
PHYIN PA DANG, ,SANGS RGYAS DE BZHIN BYANG SEMS DPA', ,MING TZAM

ZHES NI RAB TU GSUNGS, ,DI NI SGRA DON 'GOG PA STE, ,DNGOS PO SEL
BAR BYED PA MIN, ,DE BZHIN DU NI NGAG GZHAN LA'ANG, ,DON RNAMS
NGES PAR SHES PAR BYA, ,DE NYID RIG PAS MING RNAMS KUN, ,DON JI LTA
BAR YANG DAG TU, ,DMIGS PA MED NYID DE YI PHYIR, ,SGRA 'DI ZLOG PAR
BYED MA

YIN, ,RAB 'BYOR GYIS NI SGRA DANG SGRA'I, ,DON NYID GNYI GA 'GOG PA
NA, ,DE YIS BYANG CHUB SEMS DPA'I MING, ,MA MTHONG ZHES NI BRJOD
PAR 'GYUR, ,SHES RAB PHA ROL PHYIN NGAG GANG, ,DI TZAM GYIS RTOGS
MIN PA MED, ,ON KYANG 'DI DAG DE SHES PA, ,RNAMS KYIS ZHI BA

MO'I BLO YIS DPYOD, ,SKABS KYI DON NI SPONG BA DANG, ,DON GZHAN
RTOG PA GANG YIN PA, ,DE NI SHES RAB PHA ROL TU, ,PHYIN PAR GZUGS
BRNYAN YIN PAR 'DOD, ,SHES RAB PHA ROL PHYIN RTEN CAN, ,BSDUS DON
'DI TZAM NYID YIN TE, ,SLAR YANG DON GZHAN LA BR TEN NAS, ,DON DE

NYID NI BZLAS BA YIN, ,SHES RAB PHA ROL PHYIN MA NI, ,BRGYAD STONG
YANG DAG BSDUS PA YI, ,BSOD NAMS THOB GANG DES SKYE BO, ,SHES RAB
PHA ROL PHYIN GYUR CIG, '**PHAGS PA SHES RAB KYI PHA ROL TU PHYIN MA
BSDUS PA'I TSIG LE'UR BYAS PA, SLOB DPON PHYOGS KYI GLANG PO'**

**ZHAL SNGA NAS MDZAD PA RDZOGS SO,, ,KHA CHE'I PAndi TA TI LA KA KA
LA SHA DANG, LO TZ'A BA DGE SLONG BLO LDAN SHES RAB KYIS BSGYUR
BA'O,,**

[translation of the preceding portion in bold typeface:] *A Brief Presentation of the Exalted
Perfection of Wisdom, Set in Verse*, as composed by Master Dignaga, is completed.
Translated by the Kashmiri pandit Tilaka Kalasha and the [Tibetan] translator Bhikshu
Loden Sherab.

Example 2: an entry from the ACIP Master Catalog

Catalog Number: TD3809
Title (Tibetan): 'PHAGS PA SHES RAB KYI PHA ROL TU PHYIN MA BSDUS PA'I TSIG LE'UR
BYAS PA
Title (English): A Brief Presentation of the Exalted Perfection of Wisdom, Set in Verse
Title (Sanskrit): A-rya Prajn~a-pa-ramita-sam.grahaka-rika-
Skt. Title (in situ): A'ARYA PRADZNY'A P'A RA MI T'A SAM GRA HA K'A RI K'A
Author (Tibetan): PHYOGS KYI GLANG PO
Author (Sanskrit): Digna-ga
Author (English):
Author Dates:
Section/Volume: SHES PHYIN, Vol PHA
Folios/Pages: 292B-294B
Edition: SDE DGE
Published Date: 1730
Directory Path: tengyur\sherchin\sherdu
Language(s): T
Common Title(s):
End Title:
Database File Name: TD3809M.ACT
Input File Name(s):
Byte Count: 8,600
Status: M
Description:
Notes:

Example 3: an entry from the Catalog of the St. Petersburg Tibetan Collections

S:45001

N:B9409/21

T:*, ,RDO RJE 'DZIN PA CHEN PO SKAL LDAN RGYA MTSO DPAL BZANG

PO'I ZHAL SNGA NAS KYI GSUNG LAS DPAL 'KHOR LO BDE MCHOG DRIL BU

PHYI DKYIL LHA LNGA'I BDAG 'JUG DNGOS GRUB CHAR 'BEBS ZHES BYA

BA BZHUGS SO,,

K:N

E:N

X:2

B:BDE DKYIL

A:SKAL LDAN RGYA MTSO

Y:N

F:XYL

M:T

H:W,Y

G:G

R:G

V:NGA PA

P:1A-21B

Q:5

O:8.5 x 52.4

I:6.0 x 43.4

L:N

D:N

C:ZHES DPAL LDAN 'KHOR LO BDE MCHOG DRIL BU PHYI DKYIL LHA
LNGA'I DKYIL 'KHRO DU BDAG NYID ZHUGS NAS DBANG BZHI LEN PA'I
TSUL 'DI NI MKHAS GRUB THAMS CAD MKHYEN PA'I BDE MCHOG L'U AI
PA'I DKYIL CHOG BDE CHEN ROL MTSO LA GZHI BYAS NAS 'DON PA CUNG
ZAD SGYUR DGOS PA RNAMS RANG BZOS MA BSLAD PAR RGYUD RGYA GZHUNG
BOD KYI BLA MA GONG MA'I GDAMS NGAG RANG GI RTZA BA'I BLA MA'I
GSUNG GI BDUD RTZI DANG MTHUN PAR BYAS PA'O,, ,

[Monastic seal #2 -- imprinted on the text in the catalog entry given above:]



Here is a sample entry for a single semi-fictitious title that illustrates the **standard format** for the catalogs of the Tibetan collections in St. Petersburg and Ulaanbaatar:

S:00001

(Serial number of item, which is simply the order in which the titles are entered into the catalog, with no other logic.)

N:B6388/1

(Number of the title currently, in the system of the library of the Institute or Library (if any). The number after the slash is new and represents the position of the work in a single volume of multiple works. This number is being added in ink on the title page as volumes are cataloged. In the case of a loose, random page located within a volume, a lowercase alphabetical letter will be placed after the slash in order not to disturb the total number of titles in the volume. No attempt will be made to collate these random pages into their proper volume until the catalog is completed, since this will allow the collation to be performed accurately and quickly.)

T:RDO RJE GCOD PA

(Title in Tibetan as it appears on the title page, written in ACIP input code.)

K:BADZRA TZTSEDIKA

(Title in Sanskrit as it appears on the title page in Tibetanized Sanskrit. These titles are primarily ornamental and often appear with incorrect spelling, or illogically divided into syllables. The native Tibetan operators entering the data onto diskette also have trouble reading these difficult stacks correctly, and we can anticipate additional errors on their part. Nonetheless this data will be of some benefit to scholars seeking Sanskrit equivalents.)

E:LT

(Extra languages appearing either on the title page alone, or throughout the whole work [such as in bilingual dictionary in interlineal format]. Current code options are: LT [Lanycha script, title page only]; LW [Lanycha script, throughout whole work]; MT and MW [same for Mongolian]; DT and DW [same for Devanagari]. No attempt is made to input the entire extra language, but simply to indicate its presence.)

X:3

(X represents a seal or stamp that has been pressed onto the title page, a common monastic practice to identify the library or press from which the work has come. Each time a new seal is encountered, it is photocopied and added to a master chart and given a number. These seals have been scanned into computer files, and are included in our

database; we have not yet attempted to relate the seals to particular institutions from which the texts were obtained. Some title pages have multiple stamps; these will be represented by numbers separated by commas.)

B:RDO GCOD

(Brief title of the work as it appears by tradition in a small box on the end of the woodblock.)

A:SH'AKYA THUB PA

(Author of the work, only as reported in the colophon to the particular title. If the name of the author has been identified by other means it will be included in braces, for future verification.)

Y:SHING PHAG

(Year of publication or composition, only as reported in the colophon to the particular title. If either year has been identified by other means it will be included in braces, for future verification. This applies as well to attempted correspondence to the Western year.)

F:XYL

(Format of the book. Current options are **XYL** [xylograph], **MS** [handwritten manuscript], or **TYP** [metal typography].)

M:T

(Material used, meaning apparent source of paper. Current options are **T** (Tibet), **R** (Russia), **I** (India), or **C** (China). The decision of the apparent source of paper will be somewhat inexact by nature, and will be made by the Tibetan input operators and Russian experts on site.)

H:W,Y

(Hue of paper. The first letter refers to the color of the paper, which in almost every case will be white or off-white, both indicated by **W**. Works on black paper do exist at the Library and will be indicated by **B**. The letter following the comma gives the color of the traditional dye smeared on the side of the volume, which can aid in the identification of the source of a text. Current options are original **White**, or else colored **Red** or **Yellow**.)

G:G

(Grade of paper. A subjective decision by the operator about the strength and thickness of the paper. Current options are **Good**, **Medium**, or **Bad**.)

R:G

(Readability of printing. Subjective decision by operator about whether the printing is clear and easy to read, or whether for example parts are badly smudged and illegible. Current options are **Good**, **Medium**, or **Bad**.)

V:KA PA

(Volume number, according to traditional Tibetan system of assigning letters in alphabetical order.)

P:1A-15B

(Page numbers of text, as they appear in text.)

Q:6

(Quantity of lines per page, taken from the front side of the fifth folio, if there is one. Otherwise from any page typical for the majority of the text.)

O:5.2 x 20.6

(Outside dimensions of folio; the dimensions of the paper page. Again taken from the fifth folio if available, otherwise from any typical page. Expressed in centimeters to nearest tenth.)

I:4.3 x 19.4

(Inside dimensions, meaning dimensions of the printed area, normally the dimensions of the traditional line box carved around the text. Again taken from the fifth folio, front side, or from any other, typical folio if it is not available.)

L:LHA SA

(Location mentioned in the colophon to the particular text. Normally the place of printing, but sometimes place of composition of the text.)

D:1AL-RJE RIN PO CHE, 1AR-RJE BTZUN BYAMS PA

(Drawings carved or painted on the folio, with location and identification of figure copied from accompanying caption. In the above example, **1** indicates folio number; **A** front side of folio; and **L** the left side of the folio, with the caption for the figure following a hyphen. When a certain ornament covers the entire folio this is indicated in the form of **1A-RGYAN**.)

C:PAR BYANG SMON TSIG CHA TSANG,

(Normally the entire Colophon.)

Example 4: some entries from the *Great Tibetan Dictionary*

@772 {CHA}

CHA, BOD YIG GSAL BYED DRUG PA 'DI'I NGA RO 'DON TSUL LA SKYE
GNAS RKAN DANG, BYED PA LCE DBUS, NANG GI RTZOL BA MGRIN PA PHYE
BA DANG, PHYI'I RTZOL BA SROG CHEN SGRA MED,...1. 1) MING GZHI'I
MA NING YI GE ZHIG ...2) BRJOD BDE'I DBANG GIS DRANGS PA'I MING
GI CHA SHAS SHIG ...SKAD CHA,...DGOS CHA,...RGYAN CHA,...DPE
CHA,...MTSON CHA,...YIG CHA,...RAS CHA,...LAG CHA,...2....1)
ZUNG,...CHA SGRIG PA,...CHA 'GRIG PA,...CHA MIN YA MIN,...LHAM
CHA GNYIS,...GDUNG MA RING THUNG PHRA SBOM CHA MA YIN PA,...2)
RGYUS MNGA',...CHA MA MCHIS,...GAR SONG CHA MED,...SANG PHOD
THAN PA GTONG GI YIN MIN DA LTA CHA MA 'TSAL,...3) DUM BU'AM
TSAN KHAG ...BRGYA CHA,...STONG CHA,...STOD CHA,...SMAD
CHA,...THOB CHA,...CHAG CHA,...'PHAR CHA,...ZLA BA'I DKAR
CHA,...BYI LA NI GCAN GZAN GYI CHAR GTOGS PA RED,...4)
'PHRIN,...CHA SKUR BA,...PHAN TSUN GYI BAR DU NGAG THOG NAS CHA
'BUL BA,...5) GOS KYI ZHENG TSAD,...MI GZUGS STOBS CHEN PO LA
GYON PA CHA CHEN PO DGOS,...

CHA RKYEN, MTHUN RKYEN,...'TSO BA'I CHA RKYEN,...SNGON 'GRO'I
CHA RKYEN,...CHA RKYEN 'DZOMS PA,...CHA RKYEN MED PA,...CHA
RKYEN GSAR SKRUN BYED PA,...

CHA KHA, (YUL) MANG TSIG STON BYED,...'U BU CHA KHA MNYAM DU NAS

ZHING LA 'GRO,...

CHA GA ,...1) NANG DU BSGRIL BA'I GOS KYI MTHA'

@773

SNE,...GOS RNYING PA LA SNAM RAS GSAR PA'I CHA GA GTONG BA,...2)

(YUL) MIG MTHA'I DKRI GA,...MIG CHA GA CAN,...BYIS PA 'DI'I MIG

LA CHA GA YOD,...

CHA GA PA, TZIR SGRA DANG BCAS CUNG ZAD 'PHUR ZHING MCHONGS

'GROS BYED PA'I 'BU ZHIG ...

Example 5: some entries from a Sanskrit-Tibetan dictionary

@052A KRITI-KRIP'An:

[KRITI] (1) GRO GA (SHING STAG PA'I SHUN PA) (2) MKHAS PA, (3) MDZAD PA,

[KRITI:] BYED PA PO,

[KRIT'I] GOMS PA CAN,

[KRIT'INRA] BYANG BA,

[KRITYA] (1) GSHED BYED, (2) BYA BA,

[KRITYANUshTHANA] BYA BA NAN TAN DU GRUB PA,

[KRITYAm] DGOS PA,

[KRITYA:] BYA BA BYA,

[KRITY'A] BYA BA,

[KRiTRIM] BCOS BU'I RAS,

@052B [KRiTVAmMA] BCOS PA,

[KRiTV'A] BYAS NAS,

[KRiPAn] (1) BRKAM PA, (2) 'CHUMS PA,

[KRiPAnA:] (1) RAL GRI, (2) 'JUNGS PA,

[KRiPTI] RMOD, (ZHING RMOD PA)

[KRiPAN] (1) MNAR SEMS, (2) 'KHREN PA,

[KRiPALU:] BRTZE LDAN NAM BRTZE BA CAN,

[KRiP'A] (1) BRTZE BA (2) BRTZE SEMS,

[KRiP'An] CHAN GRI,

KRiP'AnA:] GRI RING,