PAN INDIAN RARE MANUSCRIPTS INITIATIVE: A PROPOSAL

Sangeetha Menon
and George M Williams

Philosophy of Science Unit
National Institute of Advanced Studies
Bangalore

NIAS Report R3-98

National Institute of Advanced Studies
Indian Institute of Science Campus
Bangalore 560 012, India

National Institute of Advanced Studies
Pan Indian Rare Manuscripts Initiative:
A Proposal

***

Sangeetha Menon
and
George M. Williams

***

The National Institute of Advanced Studies has begun a rare manuscripts digital preservation initiative. It has started modestly with a single text, recognized by the world for its philosophical and literary beauty. The *Bhagavad Gītā*, called by Swami Vivekānanda the "Gospel of Hinduism," is a universal symbol of India's great cultural heritage.

The *Bhagavad Gītā* is a pivotal text of Indian culture, philosophy and religion. It is vested with teachings of philosophical, psychological, ethical as well as of metaphysical import. Recent studies find the relevance of the *Gītā* in understanding even concepts like consciousness and time.

Some might question that anything more need be done about the *Bhagavad Gïtâ*. It is already preserved, certainly the most "preserved" Indian text—both by number of manuscripts archived and translation into more languages than any other ancient Indian text. So when so many other documents remain unknown and are often unpublished, why start a rare document initiative with the *Bhagavad Gïtâ*.

Why Start with the *Bhagavad Gītā*

The main reason for the *Gītā*'s selection arose from its very innocence. It should be the most likely manuscript that archivists and librarians would allow to be copied with a new technology. It is not a sectarian document so there are no religious grounds for refusing to allow its preservation digitally. Also there is no great fear that there is some commercial venture masquerading as scholarly research. CDs have already been produced, books written. A major critical edition was published decades ago based on several hundred manuscripts. Yet no comprehensive collection of manuscripts of the *Gītā* has been produced. Some will certainly object: "But don't you know of the work at Poona or what the Germans did in the 1930s?" Much work has been done but it does not come close to what is now possible.

The initiative began with two goals: (1) to test manuscript copying technology and (2) to digitally copy manuscripts of the *Bhagavad Gītā* for inclusion in a computer database open to the nation and the world.

The new manuscript copying technology has been called a "pioneering effort" by none other than the faculty computer services director at the University of California, Berkeley, who advises the Berkeley Multivalent Document Initiative's 30 plus field research and microfilming projects. According to Dr. Howie Lan of Berkeley, this new approach to manuscript copying and their digital archiving into computer databases may revolutionize electronic archiving efforts because of the savings in cost and human labor.

NIAS has attempted to cover expenses of the team in India. The project has been conceived as having five phases with the first phase to test the digital archiving equipment and procedures, visit a variety of institutions which possess manuscripts of the *Bhagavad Gītā*, and attempt to demonstrate, at so early a phase in the copying of existing

manuscripts of the *Bhagavad Gītā*, that there is much that has not been done before. Further it is hypothesized that there is very little time, perhaps as little as twenty to thirty years, before manuscripts available today will no longer exist in India. [See *Phases of the Bhagavad Gītā Manuscript Preservation Initiative* below]

Initial Success

All the goals of the first phase have been exceeded. The equipment and digital capture procedures worked superbly. Where access was permitted a greater range of textual variants and variables were discovered than hypothesized. There was greater cause for alarm that rare cultural treasures of India were being rapidly lost to neglect, loss, damage, etc. But there was also enough hope that this can be reversed that NIAS is calling for a much larger rare manuscripts preservation and digital archiving initiative for all of India.

The first and foremost concern was to copy enough surviving manuscripts of the *Bhagavad Gītā* into an electronic database to test the feasibility of the project and access its chances of success. That concern has been met.

Visiting Representative Archives and Libraries

Most of the archives and libraries were poorly funded and supported. The procedures at more than 75% of the institutions visited were highly protective (with some good reasons, of course), bureaucratic (requiring long forms and longer waiting periods), and often disappointing (manuscripts catalogued could not be found, manuscripts listed in the catalog as "in good condition" were highly damaged by worms or silverfish, even microfilm was often missing or already at the end of its maximum thirty year usefulness and therefore useless). [See *Appendix A*:"Toward a Modern Philosophy of Access to Information"]

Almost every institution reported that few if any Indian scholars came to or used their Sanskrit manuscripts archive. It was a rare and seemingly happy occasion when foreign scholars came. For that reason and many others, the library personnel seemed to be discouraged—by their surroundings, resources, and lack of perceived importance of their tasks. A few were clearly lacking the needed skills to read the various scripts in which Sanskrit texts of the *Bhagavad Gītā* are written. [See color insert: "Ancient Scripts"] Some lacked any training in modern, inexpensive preservation methods. And most did not have historical training and interest in maintaining the historical data about each manuscript (when written, who commissioned this copy, locale or region where written, even place where and when it was last purchased or from where it came as a gift to the library or archive, etc.).

Variants in the Text of the *Bhagavad Gītā*

So many variants in the manuscripts of the *Bhagavad Gītā* were found that the multivalent database had to be expanded. So far twenty-three textual, manuscript and scribal variables have been identified. These certainly apply to other texts than the *Gītā*, and with some standardization this project can become a model for all of India's literary and philosophical treasures. No study, certainly no dissertation, should be attempted in the future without addressing these variables when they apply to that particular text. Too many studies in the past have found a manuscript in the National Library in Delhi, translated it and completed a philosophical or thematic analysis—and stopped. But that approach ignores the range of possible variables or variants in a particular text and the issues tackled by a critical and comprehensive edition.

Primary Variables Used in the First Phase

I. Script in which Sanskrit is written

II. Dedication before beginning text

III. Type of writing material
IV. Way of beginning new chapter
V. Way of ending chapter
VI. Way of identifying the four speakers
VII. Variables because of line length (Sandhi)
VIII. Way of ending *Śloka* and marking *pāda*(s)
IX. Missing and Additional *Ślokas*
X. Variations with individual *Ślokas*
XI. Scribal mistakes (misspellings, grammar)
XII. Variances in Meter
XIII. Scribal differences in writing script
XIV. Number of different handwritings
XV. Physical characteristics of individual manuscripts

Secondary Variables

XVI. Transliterations
XVII. Translations
XVIII. History of Individual Manuscripts
XIX. Location of Individual Manuscript
XX. Publication and/or reproduction history
XXI. Glossary and Word Study
XXII. Commentaries and thematic analysis
XXIII. Manuscript Families

Other variables will evolve according to the needs of the next phases. Thus, the usual questions and arguments about dating a text, the original version of the text, the author or authors of that text, etc., begin to have an objective field of data from which to gain new insights. Authorship and dating of the writing are central concerns of critical editions that attempt to reconstruct the "original text." But meaning of various perspectives will be the central concern of a comprehensive database.

Dating of the *Bhagavad Gītā*

Though the *Bhagavad Gītā* is a part of the *Mahābhārata* (*Bhiṣma Parva*) we cannot be sure of its date, since the *Mahābhārata* includes thought of different periods. While Telang says that the *Gītā* goes back to a date earlier than the third century before the Common Era (BCE), R. G. Bhandarkar thinks that the *Gītā* is at least as old as the fourth century BCE. Garbe assigns the original *Gītā* to 200 BCE and the present form of it to 200 of the Common Era (CE). It is also interesting that the *Purāṇās* (second century CE) contain many other *Gītā*s composed in the manner of the *Bhagavad Gītā*.

Text of the *Bhagavad Gītā*

The accepted text of the *Bhagavad Gītā* possesses 700 *ślōkās*, the number confirmed by Śankarācharya. It is argued that Lord Kṛṣṇa could not have taught all the seven hundred verses to Arjuna on the battlefield. The essential verses he recited then might have been elaborated later by the narrator (Veda Vyāsa) into an extensive work. Variants of the text (with the accepted number of 700 verses) are known, most with 715 or less verses. But there is one ancient list in the Mahābhārata (6:43:4) which has 745: "Keśava spoke six hundred and twenty *ślōkās*, Arjuna fifty seven, Sanjaya sixty-seven, and Dhṛtarāṣtra one; such is the extent of the *Gītā*." Otto Schrader discusses a Kashmir

recension of the *Gītā* from which two stanzas are missing, and one is in a different place. There are nearly 250 variant readings from the accepted versions of Śankara. There has been no complete survey of all extant bark, leaf, and paper manuscripts of the *Gītā*—for obvious reasons. But today the technology exists to begin such a work.

Beginning Comprehensive Preservation

The most important thing for India to do to preserve its ancient manuscript treasures is a twofold step: (1) immediately begin a ten year project of digitally copying all ancient manuscript with new and relatively inexpensive technology and (2) leading the world in the develop and use of speech recognition technology. The reason these two things fit together is not immediately clear. So they must be unpacked.

Sanskrit and Computerization: Sanskrit, like any language, is both oral and written. Sanskrit as an oral language is one of the most regular in grammar and pronunciation. It, in fact, is so rigid that it has been called "artificial"—great for finished literature but poor for conservation. Be that as it may, oral Sanskrit is perfect for computerization. The literal-mindedness of the computer has to be stretched less with so regular and perfect a language.

Written Sanskrit has a clear and precise relationship to Oral Sanskrit and consequently is also a perfect candidate for computerization. Already projects around India have begun this process. But other priorities have actually thwarted Sanskrit's computerization. Some are clearly political. Therefore, it must be stated again that it is Sanskrit's regularity or "artificiality" that makes it the best language in the world for computerization—both written and oral language systems. No living language, even those derived from Sanskrit, meets these criteria—their irregularity, flexibility and colloquial inventiveness are precisely at fault.

Written Sanskrit can be represented with any script—but since the Roman script or alphabet has only 26 letters, this representation is incomplete and less than adequate. [Alphabetic, written languages also have important differences from syllabic, written languages. This need not concern us at this stage but might suggest another type of keyboard or data entry system instead of just copying the Qwerty keyboard.] Diacritical marks are added to make up for this deficiency in a Roman alphabet representing the Sanskrit syllabry. This should be standardized in 2-bit code, also recognizing the differences between alphabetic and syllabic languages. Using a standardized, universal 2-bit code would allow instant switching by computer from any world script (Roman to Devanagari to Telegu) to any other script.

From Royal Manuscripts to Family Treasures

Until perhaps the fifth century of this era *pandits* and priests spoke or quoted Sanskrit from memory and continued to resist different writing systems that were already being employed. Rulers probably commissioned some of the first manuscripts—and some of the rarest have golden margins and occasional gold lettering. Gradually manuscripts were used for study and later even objects of veneration in the temples themselves. Eventually Brahmin families in villages would employ a scribe and care for and maintain their own copies of the *Bhagavad Gītā* —among other texts.

Therefore, an extraordinary range of manuscripts will be found, digitally copied, classified and studied. [See color insert: *Ancient Scripts*]

Phases of the *Bhagavad Gītā* Manuscript Preservation Initiative

I. Phase One: Testing and Feasibility

The initial phase involved testing a new technology in as many types of institutions (libraries, research centers, private and religious collections). The goal was to digitize and get into the computer as many different types of manuscripts of the *Bhagavad Gītā* as possible within the time and financial constraints. Examples of eight ancient scripts were digitally preserved from three different types of ancient manuscripts (palm leaf, bark and handmade paper). Twenty-three variables were noted as much as possible about the manuscript itself, plus field notes about contact persons and archives.

II. Phase Two: Creating a Multivalent Database at NIAS and Inputting Digital Versions of manuscripts of the *Bhagavad Gītā*

This phase has been already started from the second data went into the computer and was preserved there digitally. Unlike microfilm with its short lifetime digital data can be preserved on CD-ROM for about a century—and with reasonable maintenance permanently.

The database will have the following characteristics in order to share India's national treasure with the world. It will be copyrighted to protect against commercial exploitation, yet free for educational usage. It will be open to all via the Internet. There will be an interactive bulletin board for scholars around the world to interact with scholars in India as the database expands. The database will be expandable and extensible.

III. Accelerated Data Collection and National Digital Archive

During this phase librarians and archivists will be brought to NIAS and trained in the new archival technology. The workshops will involve training in the use of digital equipment and computer input. One aspect of the workshop will be a review of the latest and best inexpensive preservation methods for their manuscripts. The syllabus for the course will include, among many other things, examples of different ancient scripts in which ancient *Bhagavad Gītā* manuscripts were written. A section of historical notation concerning each institution's particular manuscripts will be taught.

Probably by working with one manufacturer a set of thirty digital computer input devices will be obtained. (This phrasing is purposefully vague to allow for negotiations with companies for a major donation. This project can be replicated for libraries and archives worldwide. Any company having their equipment used in India would receive sufficient benefit to consider a major donation of needed equipment.) Even if no equipment is given, the cost of thirty loaners will not be prohibitive.

Each person successfully completing the NIAS course will be loaned a set of "digitizing tools" and sent back to the library/archive to digitally copy all of each institutions' manuscripts on the *Bhagavad Gītā*. Then they will, preferably, personally return the equipment to NIAS so that the raw data can be put into the computer. A follow-up seminar would encourage the participants toward returning to the library and beginning a national initiative to preserve digitally all of India's rare manuscripts. Institutions would be encouraged to change to this new and permanent digital technology and to purchase the needed equipment through a combined discounted, purchasing plan.

New groups of manuscript librarians will be brought to NIAS or to future regional training centers and trained, etc., as the process is repeated again and again.

Phase Three will continue until there is no further extant manuscripts of the *Bhagavad Gītā* to enter into the database and no more manuscript librarians to be trained in the new technology.

Phase IV. Alphanumeric Conversion and Speech Recognition

Phase Four should begin immediately as funds are available. Preserving the manuscripts as pictures or graphics is a good beginning. But it is not enough. For the database to be a knowledge and educational resource the graphics must be recognized as computer-readable/recognizable script. As mentioned before the Oral Language of Sanskrit can be written in many scripts. If scripts are utilized that are standardized (and India may wish to revise the code as a version two of Unicode World Script), then the scripts in which Sanskrit are represented will be interchangeable instantly by the computer.

As mentioned earlier in this paper, Oral Sanskrit is the best language in the world for speech recognition and its input into the computer. So all efforts must be brought together in a national seminar and progress noted. At that point it will become clear just how to proceed in making India the world leader in speech recognition technology—because Sanskrit is the most suitable language in the world for this application. [Latin is a poor second choice.] Once Oral Sanskrit is computerized that success can be extended to other South Asian languages.

Instead of retyping each manuscript into the computer, or only typing variants found in a particular manuscript into the computer, manuscripts in any ancient script can be read by some of the few old scholars who can do these readings consistently enough for easy computer recognition. The computer will transliterate the Sanskrit into any and every font/scripting system in the entire Unicode system of language scripts. It is that direct, once speech recognition is solved.

V. Phase Five: Creating a Living, Interactive, Comprehensive Edition of the *Bhagavad Gītā*

The first comparisons between versions of the *Bhagavad Gītā* have already been started. However, in any real sense the process of utilizing a comprehensive database of all versions of the *Gītā* will take at least a decade of collection, scholarly study and interaction, and reflection before the fruits of this process will become transparent. This has the chance of becoming the first comprehensive interactive database of any ancient manuscript (including the Christian Bible, etc.). The difference between a critical edition of an ancient text based on a few hundred manuscripts and a comprehensive interactive database cannot be over-emphasized. The latter will be available to all via the Internet. [See *Appendix B*: "A Comprehensive, Interactive Database..."]

By this time of the project most manuscripts of the *Bhagavad Gītā* will be seen as rather ordinary (and there will be thousands of these); but some rare manuscripts may prove to be older than the best preserved *bhāṣyās* of the great commentators. What that means is that the database will be open to scholarly discussion and interpretation. This can be facilitated by electronic bulletin boards for scholarly, and even personal, discussions.

India's great cultural heritage will be preserved for discussion and not lost. That is the whole point of this initiative.

Appendix A:
Toward a Modern Philosophy of Access to Information

Every scholar and unfortunately many students as well have run into libraries and archives run by the 3H's: Hide, Hoard, and Harass. A national archive of *Bhagavad Gītā* manuscripts would try to reverse the following scenario.

One asks for a particular manuscript and you are told it doesn't exist. Then when you produce an accession number that another scholar has cited of their manuscript's existence, some progress might be made. But even then, you can be told to come back tomorrow, next week, or even next month, because the person responsible is "out of station." If you are unlucky enough to have no accession numbers for a particular archive, worst things can happen. You may be told that there is no list of manuscripts, or no published list that the public can study, or only an incomplete list.

Hoarding is another problem. Some public libraries have no photo-reproduction [xeroxing] facilities and may only allow you to take handwritten notes. This practice is so anti-scholarly that it was quite mercifully found only in five per cent of institutions visited.

Harassment is a serious charge and it certainly reflects the bias of scholars who think that the norm should be cooperation and reasonable costs. But there still exists a "spy mentality" among some that suspect any request to use a manuscript. Questions vary from "what are you going to do with this data?" to "why do you want to see the entire manuscript?" It is not appropriate to say simply that you wish to learn whatever can be learned from it. Forms are so lengthy and complex that they can require most of a day to complete and sometimes more time than it takes to read the material requested. Often a non-librarian, usually a political appointee, has the last word and only after their signature is acquired do you even get a look at the manuscript. Heaven help you if accession numbers are inaccurate! In several instances we are still waiting for the officials up the chain of command to decide if our request for photo-reproductions of their *Bhagavad Gītā* manuscripts are legitimate. Then there will be a much longer process about computing the costs, getting payment to them in advance or, worst, being required to return to a distant archive to stand there while the copying is being done.

If the 3H mentality is to be changed, there needs to be a clear understanding that public acçess to information is in the best interest of a culture and a democracy. The 3H mentality is an artifact of governance and management styles that are no longer appropriate in the Information Age.

One of the hoped for outcomes of a national initiative to preserve India's manuscript treasures is a new emphasis on proper use and sharing of these treasures with the world.

Appendix B:
A Comprehensive, Interactive Database of all Existing Manuscripts of the *Bhagavad Gītā*

"What will you do that hasn't already been done? Do you think that you will be able to produce a better critical edition [than the one edited by Prof. S. K. Belvalkar]?"

These two questions capture the crux of the problem. First, we do not intend to produce a critical edition as it has been conceived by Western text criticism. That kind of enterprise entails a presupposition which we think is invalid for most Indian texts--

namely, that a critical edition will produce the original version as it was written or given the very first time. Second, scholars recognize that most Indian sacred texts have evolved through an oral tradition in families or *Śākas*. There is, of course, another notion that *śruti* was either uncreated and simply heard or created for a sage to hear and remember. But even this notion speaks of multiple tellings, so that no single account would be comprehensive. In the case of the *Bhagavad Gītā*, there would be Kṛṣṇa's conversation with Arjuna (perhaps represented by chapters two and three, quickly summarizing everything under battlefield conditions). There is also Sanjaya's version spoken to Dhṛtarāṣṭra. There is Vyāsa's original 745 *ślōkās* (if that tradition is the correct number). And there is also an account that Ganeśa wrote down as the scribe of Vyāsa, perhaps accounting why there are only 700 *ślōkās* now. Why all this is mentioned is to point out that there should be multiple versions of the *Bhagavad Gītā* of varying lengths.

If a critical edition has as its purpose to reduce a text to one and only one "original edition," then you should ask where is this agenda coming from. Indian hermeneutics have not thought that a reduction of the richness of variants to only one account was sufficient or truthful.

This reductive approach only works for a writing with a single author or writings that are supposed to have one divine author who dictated everything verbatim to a single scribe. Although it was once believed that the *Christian Bible* came into being by a "word-for-word" dictation of an "original text," that notion has been abandoned long ago as fanciful, unworthy of the facts, and even harmful for a full understanding of Christian teachings. But this method of producing critical editions of sacred texts has continued and was used on the *Bhagavad Gītā*, even though it is clearly inappropriate.

An adequate approach which would honor the richness of the Indian cultural and philosophical heritage would be a comprehensive edition--critically examining all existing texts but not reducing them to a single version. The approach will have to be inclusive--including all versions and variants.

So a response can now be made to the question about "hasn't everything already been done?" and "are you able to produce a better critical edition?" The S. K. Belvalkar critical edition of 1946 (reprinted in 1968) based its findings on the number of manuscripts thought to exist at that time--less than 200! Even then, there is only a claim that about 125 manuscripts were used.

Belvalkar's *Editorial Notes to the Reprint of the Bhagavad Gītā* stated: "From the descriptive catalogues, published reports, and the lists of MSS. available in public libraries in India, as also from a special search instituted in private libraries in various part of India, the existence of more than 125 MSS. of the Bhīṣmaparvan (text only) has been so far established, about half of them being written in Devanāgari characters, about 20 in Bengali characters, some 15 each in Grantha and Telugu characters, over 10 in Malayalam characters, but only 1 in Śāradā characters." [p.ix] One can now visit three or four libraries or archives and see more manuscripts of the *Bhagavad Gītā* than that.

A comprehensive edition will continue to grow with the intention of one day having a photographic image and digital version of all surviving manuscripts of the *Bhagavad Gītā*. Trying to produce an original version of the *Bhagavad Gītā* can be abandoned and replaced by trying to produce a comprehensive edition. The fullness of the Indian heritage will be honored when all families and branches of transmission and interpretation are preserved and studied. Finally meaning of ideas becomes the central focus in all its variety.

Acknowledgement

Many thanks to the Chief Advisor of this project, Prof. B. V. Sreekantan (Dr. S. Radhakrishnan Chair, NIAS), and to Dr. R. Narasimha, the Director of NIAS.

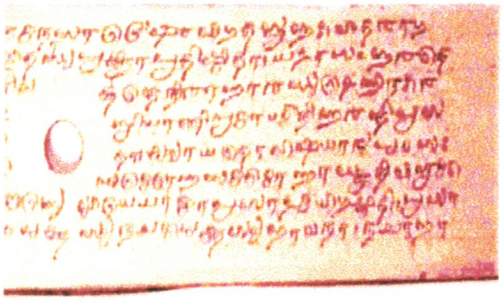
Thanks also to Prof. Lakshmithathachar of Academy of Sanskrit Research, Melkote, for a catalogue of catalogues of manuscript libraries.

Further appreciation to all the librarians and archivists who enthusiastically helped us.

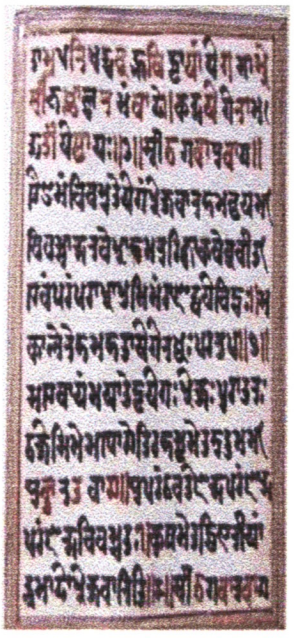
Dr. Sangeetha Menon completed her doctoral thesis on the "Concept of Consciousness in the *Bhagavad Gītā*," soon to be published. She is a faculty member of Consciousness Studies in the Unit of the Philosophy of Science, NIAS.

Prof. George M. Williams has been teaching the *Bhagavad Gītā* to American students for the past thirty years in his Asian religions' courses. He is a member of the Electronic Publishing Committee of the American Academy of Religion (AAR). He has acted a project manager for CD ROM publication for the AAR—a voluntary and non-remunerative position. He has pioneered digital archiving techniques and Internet scholarly publication and brought some new and previously untested technology to NIAS for this project. Prof. Williams is a Visiting Professor (Sir Ashutosh Mukerjee Chair) at NIAS.

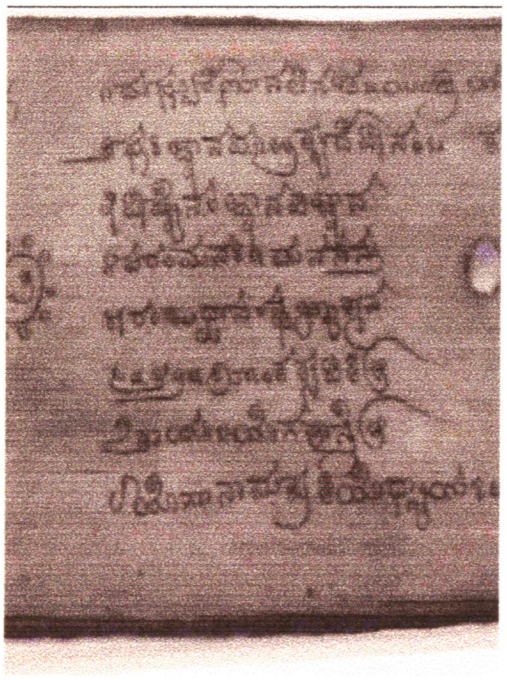
Selections from the *Bhagavad Gītā:* Palm Leaf & Paper MSS.



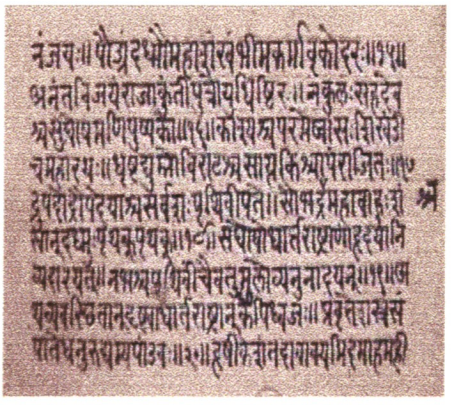
Grantha Script. **Palm Leaf.**
Academy of Sanskrit Research, Melkote



Śārada Script. **Handmade Paper.**
Adayar Library, Madras



Kannada Script. **Palm Leaf.**
Academy of Sanskrit Research, Melkote



Devanāgari Script. **Handmade Paper.**
Osmania University Library, Hyderabad