



An Introduction to the Textual History Tool

Diptesh Kanojia, Malhar Kulkarni, Pushpak Bhattacharyya, Sayali Ghodekar,
Irawati Kulkarni, Nilesh Joshi, and Eivind Kahrs

IITB-Monash Research Academy

IIT Bombay

Monash University

Cambridge University

diptesh@cse.iitb.ac.in

Contribution

Our contribution is the construction of a tool which can:

- Enable philologists to enter data for any text (including vulgate data) and various manuscripts versions.
- Compare the different versions in a separate mode.
- Edit the data while viewing the manuscript labels.
- Constructs phylogenetic trees based on distance matrices which can be constructed based on different approaches.
- Demonstration of the historical evolution of a text in one-go.

Roadmap

- Introduction
- Tool Development
- Methodologies for Tree Construction
- Discussion and Conclusion

Introduction

Text as a historical document

Texts are primary sources of History-Physical + Intellectual- of any land.

Texts themselves have a history and evolve in the course of time.

Intellectual traditions get built around texts, ex: Text of Panini's grammar and the Paninian grammatical tradition that continues till date

Texts in the Indian tradition show stages of evolution, important in the reconstruction of history of that tradition

Can digital technology help study/record the evolution of texts in Indian tradition?

Textual History and Criticism

Written transmission is carried out through copies of the text, also known as manuscripts. Historically, manuscripts were written or copied by one or more scribes.

Transmission of the text from one source to another generates variants which differ significantly when compared to each other.

Previously, philologists used paper-based methods for various purposes viz. collation, description of manuscripts, inter-relation of manuscripts, apparatus creation *etc.*

We can now have tools with us, that can facilitate the data entry, storage and display of the aforementioned functions, **all on the same interface**, with a standard input mechanism.

Motivation

When we study the texts in the Indian grammatical tradition, that too, the paninian one, traditional commentators like Madhava and Bhattoji Dikshita etc., and modern scholars like Kielhorn (1887) and Kulkarni (2012a) observe that the text of the Aṣṭādhyāyī (AST) has evolved in the course of time.

The text of the sutras that Patanjali had in front of him is not the same as we have it today. It becomes important for scholars interested in the development of an intellectual tradition of linguistic thought in India to study the evolution of Kāśikāvṛtti (KV) through various sources like commentaries and manuscripts.

In order to study this stage of evolution further, when we turn to the printed text of the KV as available to us through more than ten editions, as of now, we notice that the printed editions do not present a picture of a uniform text and rather suggest that this text of the KV that we have with us today, must have evolved in a different manner historically.

Functional Units

The text of KV, as mentioned above, can be, generally, divided into its functional parts.

There are two basic divisions in the text of KV, one that of the sūtra and other of the KV. Within the KV, the text can further be divided according to its functional properties based on the type of sūtra it is commenting upon.

We present below the functional divisions in the KV on the saṁjñā sūtra.

saṁjñā: this type of sūtra introduces a technical term, and hence the KV on this sūtra is observed to contain the following functional parts:

1. Introduction of the words in the sūtra and meaning of the sūtra.
2. Examples.
3. Mention of other sūtras in which this technical term appears.

Example

1.1.1.	Sutra	वृद्धिरादैच् (१॥१॥१)
1.1.1.1	Introduction & Meaning	वृद्धिशब्दः संज्ञात्वेन विधीयते प्रत्येकमादैचां वर्णानां सामान्येन तद्भावितानामतद्भावितानां च। तपरकरणमैजर्थं तादपि परस्तपर इति खट्वैरकादिषु त्रिमात्रचतुर्मात्रप्रसङ्गनिवृत्तये।
1.1.1.2	Examples	आश्वलायनः। ऐतिकायनः। औपगतः। औपमन्यवः। शालीयः। मालीयः।
1.1.1.3	Other Occurrences of the term	वृद्धिप्रदेशाः। सिचि वृद्धिः परस्मैपदेषु इत्येवमादयः॥

Key Question

Can we build a comprehensive tool for visualizing the transmission and history of any text using the associated text information?

A tool which can visualize the multiple versions of the same text which also allows data entry for manuscript versions and thus, helping one compare these versions with each other and aids one in adapting them to a graphical model *viz.* a phylogenetic tree.

- (i) Visualize the data from earlier texts.
- (ii) Visualize the data from testimonia, and printed editions.
- (iii) Visualize the data from commentaries, sub-commentaries and abridged versions.

Tool Development

Technical Development

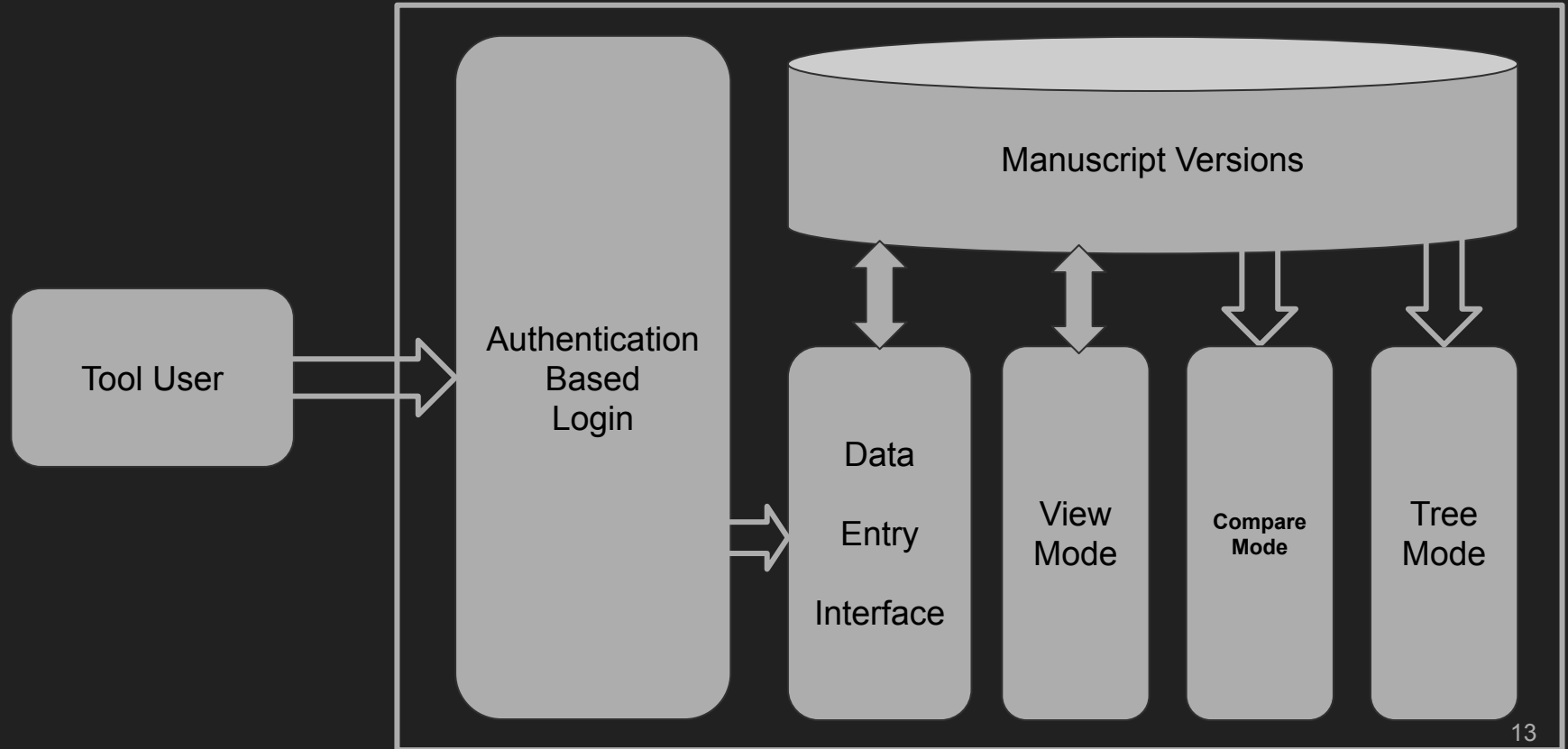
We develop the tool by using a simple server-side scripting language i.e., PHP and use MySQL for database storage on the backend. We use the Apache web server to host our tool.

Our tool allows a user to register themselves for using the tool via a simple registration form. The registration needs to be validated by the tool administrator.

The user can then logon to the tool and use the functionalities described further in our presentation and the paper.

The tool also uses Javascript, AJAX, CSS and HTML wherever necessary.

Tool Architecture



Data Entry Mode

The tool allows one to add the text in the data entry mode. It displays the vulgate text by default given the philologist has entered the vulgate text data earlier.

If not, the tool allows one to enter the data for a vulgate. In a rare case when the vulgate data is absent, the tool allows one to prepare the various versions in this mode.

These can be compared later and a phylogenetic tree can be produced based on the entered text versions which can lead to the preparation of a possible archetype.

This can help philologists prepare a critical edition of that text.

Data Entry Mode - Screenshot

Textual history analysis tool
(ग्रन्थेतिहास-दर्शक-यन्त्रम्)

Data Entry Mode Compare Mode View Mode Phylogenetic Tree Mode Logout

MANUSCRIPT	kashikavritti
LABEL ID	bh1
SUTRA ID	1.1.3. <small>WARNING: The sutra already exists.</small>

EDIT THIS SUTRA

Text (Vulgate to be shown by default)

इको गुणवृद्धी (१११॥३)

FUNCTIONAL UNITS

परिभाषेयं स्थानिनियमार्थं अनियं scibe addमप्रसङ्गे नियमो विधीयते। Add More

Historical Antecedents: Earlier Texts

The purpose of this is to provide the reader with the historical view of the text. After clicking on the earlier texts button, the user is provided with an option to choose between “Paninian” and “Non-Paninian” texts.

By choosing the option to view “Paninian” texts, the interface shows the earlier texts in the Paninian tradition, in this context, the Vyakarana Mahabhashya (VMbh).

When viewing “Non-Paninian” texts, the interface shows the earlier texts in the Non-Paninian traditions namely Katantra, Chaandra, etc.

Commentaries

Our tool allows one to enter the commentaries on the text based on the philologists' expertise.

- The text of the commentary is entered in correspondence with the functional unit of the main text. As many functional units, so many boxes.

Apart from the functional unit module, the tool shows extra entry points for any commentary instances which may exist for the text.

We allow one to enter different commentaries based on the text distribution (sutras, proses, or verses).

The commentaries can be taken into account later when building phylogenetic trees using direct and indirect evidences from them.

Direct Evidence from the Commentaries

Kulkarni and Kahrs (2019a) show that the textual evidence available in these two commentaries - Nyasa (Ny) and Padamanjari (Pm), can be classified under two broad categories:

- Direct Evidence
- Indirect Evidence

Kulkarni (2002c) mentions the effort on the part of its author to collect information from the Ny and the Pm, which can act as an evidence to reconstruct the text of the KV. Kulkarni and Kahrs (2019b) enlist the variants of the text of the KV as found in the Pm through more than 300 quotations.

Indirect Evidence from the Commentaries

Indirect commentaries are the commentaries on the direct commentaries.

Tantrapradipa (Tp) is a commentary on Ny. Therefore, it becomes an indirect commentary on the KV.

While Direct evidence is clearly visible in the text of the Ny and Pm, indirect evidence can be further classified under two categories:

- paroksha
- atiparoksha

From the given text, a portion can be highlighted based on whether it is direct or indirect. Further classification of indirect commentaries is provided in our paper.

Commentaries Entry - Screenshot

FUNCTIONAL UNITS

शि इत्येतत्सर्वनामस्थानसंज्ञं भवति, किमिदं शीतिः जसु शसोरादेशः।	<input type="button" value="Add More"/>
कुण्डानि पश्य, दधीनि। मधूनि। व्रपूणि। जतूनि।	<input type="button" value="x"/>
सर्वनामस्थानप्रदेशाः। सर्वनामस्थाने चासंबुद्धौ इत्येवमादयः।।	<input type="button" value="x"/>
NULL	<input type="button" value="x"/>
NULL	<input type="button" value="x"/>
NULL	<input type="button" value="x"/>
NULL	<input type="button" value="x"/>

COMMENTARIES

COMPLETED!

Testimonia

The text of the KV is quoted in the later texts grammatical as well as non-grammatical.

Kulkarni(2002c) collected and arranged chronologically more than 1000 such quotations as available from the later paninian grammatical tradition.

Kulkarni (2002b) studied one quotation of the KV as found in the Shabdkaustubha and showed the interrelation of KV manuscripts and Shabdkaustubha.

Our tool displays all these quotations for the sūtra under study.

Printed Editions

When “Printed Editions” is clicked, the interface displays all the printed editions’ text of the sūtra. This historical development in the printed editions is also presented in the text visualizer.

It is hoped that the amount of variation available in the printed editions will serve as a basis to understand the manuscript variants.

“Even since Kulkarni (2000) traced the manuscript sources of the edition of the KV printed in 1876, the KV has been printed more than ten times.”

View Mode

The view mode is an additional editable mode which allows the philologists to view a particular portion of the text and edit it for any required changes.

This mode does not allow one to edit the commentaries as it was made keeping in mind that a philologist might need to make small changes in the variant of a text and should not have to go through the complete data entry mode.

This also saves the philologists' time and bandwidth used for making smaller changes in the text variant.

View Mode - Screenshot

(ग्रन्थातहास=दशक=यन्त्रम्)

Data Entry Mode Compare Mode View Mode Phylogenetic Tree Mode Logout

Select from Manuscripts Select from Labels Select from Sutras

kashikavritti A6 1.1.1.

Manuscript_Text

वृद्धिरदिच् | (१॥१॥१)

FUNCTIONAL UNITS

Functional Unit 1	Functional Unit 2	Functional Unit 3	Functional Unit 4
वृद्धिशब्दः संज्ञालेन विधीयते प्रत्येकमादेयां वर्णानां सामान्येन तद्भववितानामतद्भाववितानां च तपरकरण \lacuna {9} रस्तपर इति वृष्टका वृक्षेडका खट्टाएडका खट्टेडका वृक्षेडका दियु \eyeskip {} त्रिमात्रचतुर्मात्रनिवृत्तये	आश्रुतायनः ऐतिकायनः औपगवः \lacuna {4} मालीयः \illegible {२} कुलं	वृद्धिप्रदेशाः वृद्धिर्यस्याचामादिस्तद्वृद्धम् सिचि वृद्धिः परस्मैपदेषु इत्येवमादयः	NULL
Functional Unit 5	Functional Unit 6	Functional Unit 7	Functional Unit 8
NULL	NULL	NULL	NULL

Compare Mode

In the compare mode, philologists can choose to compare two to three different variants of the text.

The variant comparison can give a philologist a better idea in a case where a standard critically edited edition is not present and a vulgate is also absent.

The philologist can compare multiple versions and come up with their critically edited version of the text.

Compare Mode - Screenshot

Data Entry Mode | Compare Mode | View Mode | Phylogenetic Tree Mode | Logout

Select First Manuscript: kashikavritti

Select Second Manuscript: kashikavritti

Select First Label: bh1

Select Second Label: ba1

Select First Sutra: 1.1.1.

Select Second Sutra: 1.1.1.

First Manuscript: वृद्धिरदिच | (१॥१॥१)

Second Manuscript: \ignoredtext(गति: ओम्) वृद्धिरदिच | (१॥१॥१)

Functional Unit 1: वृद्धिशब्दः संज्ञालेन विधीयते प्रत्येकमादिचा वर्णानां सामान्येन तद्भावितानामतद्भावितानां च । तपरकरणमिजर्थं तादापि परस्पर इति खट्टिकदिषु त्रिमात्रचतुर्मात्रप्रसङ्गनिवृत्त्ये ।

Functional Unit 1: वृद्धिशब्दः संज्ञालेन विधीयते प्रत्येकमादिचा वर्णानां सामान्येन तद्भावितानामतद्भावितानां च । तपरकरणमिजर्थं तादापि परस्पर इति खट्टिकदिषु त्रिमात्रचतुर्मात्रचोः eyeship } त्रिमात्रप्रसङ्गनिवृत्त्ये ।

Functional Unit 2: आक्षराम्नः | before ऐतिकामिनः after ऐतिकामिनः | औपगदः | औपमन्यदः | शालीयः | मालीयः |

Functional Unit 2: \laena (१) यनः | औपगदः | औपमन्यदः | शालीयः | मालीयः | ₹ :illegible (१) लः | नी :illegible (१) र्तः यः |

Functional Unit 3: वृद्धिप्रदेशाः | सिचि वृद्धिः परस्मिपदेषु इत्येवमादयः |

Functional Unit 3: वृद्धिप्रदेशाः | \folio-change (४अ) सिचि वृद्धिः परस्मिपदेषु इत्येवमादयः |

Functional Unit 4: NULL

Functional Unit 4: NULL

Functional Unit 5: NULL

Functional Unit 5: NULL

Tree Construction Mode

In this mode, our tool utilizes our work from a different paper presented at this forum.

The tool provides one with options to select the variants which need to be used to create the tree along with a select all option.

The tool allows one to create distance matrices first which form the basis of tree construction methodologies present in the tool which are shown on the interface.

Then, the philologist can go ahead create the tree using any of the methods available with us (Neighbor-joining or UPGMA).

Tree Construction Mode - Screenshot

Distance Matrix Construction

The screenshot displays the 'Phylogenetic Tree Mode' interface. At the top, there is a header in Marathi: 'दृष्टावही होइतारुप क्वापुडोइ तठो (ग्रन्थेतिहास-दर्शक-यन्त्रम्)'. Below the header is a navigation bar with options: 'Data Entry Mode', 'Compare Mode', 'View Mode', 'Phylogenetic Tree Mode' (selected), and 'Logout'. A dropdown menu shows the selected taxa: 'A10, bh6, bh7, bh8, g1, g3, Ld0'. Two buttons are present: 'CREATE DM (Using OSA)' and 'CREATE DM (Using Embeddings)'. The 'Distance Matrix' section shows a 7x7 matrix for taxa: ['bh8', 'Ld0', 'g3', 'r1', 'bh7', 'bh6', 'A10', 'g1'].

दृष्टावही होइतारुप क्वापुडोइ तठो
(ग्रन्थेतिहास-दर्शक-यन्त्रम्)

Data Entry Mode Compare Mode View Mode **Phylogenetic Tree Mode** Logout

A10, bh6, bh7, bh8, g1, g3, Ld0

CREATE DM (Using OSA) CREATE DM (Using Embeddings)

Distance Matrix

```
['bh8', 'Ld0', 'g3', 'r1', 'bh7', 'bh6', 'A10', 'g1']  
[0.0, 0.1234, 0.3247, 0.1881, 0.1875, 0.2684, 0.3717, 0.1477]  
[0.1234, 0.0, 0.3213, 0.1799, 0.1632, 0.2554, 0.3639, 0.1365]  
[0.3247, 0.3213, 0.0, 0.3111, 0.3295, 0.3534, 0.5, 0.3309]  
[0.1881, 0.1799, 0.3111, 0.0, 0.221, 0.2237, 0.4051, 0.1934]  
[0.1875, 0.1632, 0.3295, 0.221, 0.0, 0.2857, 0.3899, 0.1879]  
[0.2684, 0.2554, 0.3534, 0.2237, 0.2857, 0.0, 0.466, 0.2653]  
[0.3717, 0.3639, 0.5, 0.4051, 0.3899, 0.466, 0.0, 0.3891]  
[0.1477, 0.1365, 0.3309, 0.1934, 0.1879, 0.2653, 0.3891, 0.0]
```


Discussion and Conclusion

Observations

We do observe that there will be commentaries over commentaries of the text and we need to further inculcate the option of addition of sub-commentaries within the tool.

Please take note that in the current form, the tool can be used to successfully build a text data along with its variants and to construct phylogenetics trees.

We shall add more methodologies of tree construction to the tool and the tool is scalable in this regard.

Conclusion

We present a comprehensive data entry mechanism which can help digitize historically written texts.

This tool can be used to add multiple variants of the same manuscripts along with the commentaries based on them.

The tool can be used by a philologist to compare the various versions of a manuscript and construct phylogenetic trees based on two different methods.

The tool is scalable and we aim to add more functionalities to the tool later.

Future Work

We aim to add more methods for tree construction within the tool including the feature which can take into account the presence of a time frame for the variant.

We also wish to add a feature where sub-commentaries can also be added to the tool.

We would also like to add a different text apart from KV which can test the generality of our tool and help us create more modules which may be needed in the expansion of this tool.

We also aim to inculcate various factors like commentaries, testimonia and earlier texts to build a more accurate phylogenetic tree for a given text.

We also aim to create a faster method for tree generation (current distance matrix calculation takes a little more time than we envision for a web based service).

Tool Demonstration

<http://www.cfilt.iitb.ac.in/that>

Acknowledgements

Noemie Verdon, University of Lausanne, Switzerland.

University of Cambridge, United Kingdom.

IIT Bombay, India.

British Academy, United Kingdom.

Rashtriya Sanskrit Sansthan, New Delhi, India.

All the manuscript libraries.

Thanks

Questions?