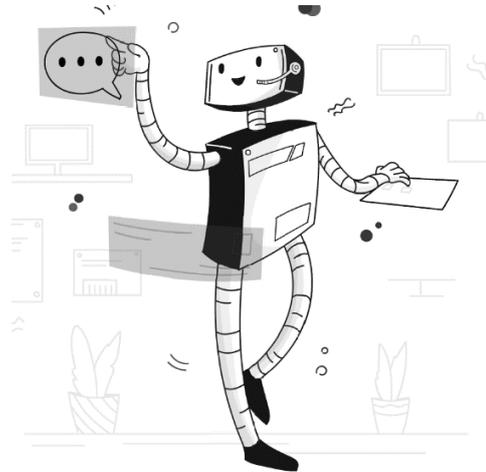




## Silvin Kořak Memorial Symposium

# Ancient Studies in the Digital Era

December 11–12, 2023



Department of Comparative and General Linguistics

Faculty of Arts, University of Ljubljana



## Conference Venue

Faculty of Arts, University of Ljubljana, Blue Chamber, Aškerčeva 2, 1000 Ljubljana, Slovenia

## Zoom-Link

<https://uni-lj-si.zoom.us/j/98206568135?pwd=VoZrazFyZjN4MjVkNopaeDJZblB6Zzo9>

Meeting ID: 982 0656 8135

Passcode: 827807

Time Zone: CET (GMT +1)

## Program

### Mon, December 11, 2023

18:00–18:30 CET

#### Welcome & Opening

**Chikako Shigemori Bučar**, President of the Linguistic Circle of the Faculty of Arts, University of Ljubljana

**Marina Zorman**, Department of General and Comparative Linguistics, Faculty of Arts, University of Ljubljana

**Mojca Schlamberger Brezar**, Dean of the Faculty of Arts, University of Ljubljana

**Daniel Schwemer**, Chair of the Ancient Near Eastern Studies, Department of Ancient Studies, Würzburg University, Chair of the German Oriental Society, Director of the Hethitologie-Portal Mainz (Remote opening address)

18:30–19:30 CET

#### Opening keynote (Chair: Chikako Shigemori Bučar)

**Gerfrid G.W. Müller** (Würzburg University), *From Filecards to AI. Development and Prospects of Hittitology from a Technical Perspective*

### Tue, December 12, 2023

9:30–10:30 CET

#### Session 1 (Chair: Gerfrid G.W. Müller)

**Metka Furlan & Luka Brenko** (University of Ljubljana), *Tribute to Silvin Košak*

**Shai Gordin** (Ariel University, Open University of Israel), *Ancient Language Processing: A Transformative Framework for the Computational Study of Ancient Texts* (Remote talk)

10:30–10:50 CET

**Coffee Break**

10:50–12:20 CET

**Session 2 (Chair: Luka Repanšek)**

**Enrique Jiménez** (Ludwig-Maximilians-Universität München) *The Electronic Babylonian Library: Taming the Cuneiform Jungle* (Remote talk)

**Preeti Shukla** (Central Sanskrit University Delhi), **Amba Kulkarni** (University of Hyderabad), and **Devanand Shukl** (Central Sanskrit University Delhi) *Revival of Ancient Sanskrit Teaching Methods using Computational Platforms* (Remote talk)

**A tour of Silvin's personal library** (Ana Mehle & Luka Brenko)

12:20–14:00 CET

**Lunch Break**

14:00–15:00 CET

**Closing keynote (Chair: Hannes Fellner)**

**Adam Anderson** (UC Berkeley) *Leveraging Quantitative Bibliometrics in Assyriology: Exploring Distant Reading, Computational Linguistics, and the FactGrid Cuneiform Wikibase Project* (Remote talk)

15:00–15:15 CET

**Coffee Break**

15:15–16:30 CET

**Session 2 (Chair: Simona Klemenčič)**

**Hannes Fellner** (University of Vienna) *Digital Philology on the Ancient Silk Road*

**Matej Hriberšek** (University of Ljubljana) *Classical Languages at the Crossroads of Tradition and Artificial Intelligence*

**Closing remarks and a screening of the theatrical performance of CTH 321, *The Myth of Illuyanka***

**The symposium will be accompanied by an exhibition on Hittite scholarship conducted in Slovenia and the place of Silvin Košak in it and beyond.**

## Abstracts

### *Leveraging Quantitative Bibliometrics in Assyriology: Exploring Distant Reading, Computational Linguistics, and the FactGrid Cuneiform Wikibase Project*

Adam Anderson (University of California, Berkeley)

Delving into the vast digital repository of Assyriological resources, comprising 150,000 books, journals, and datasets, this keynote address embarks on an exploration of methodological considerations. How can we effectively harness the potential of such a rich repository, and what novel insights can be gleaned through 'distant reading' approaches facilitated by machine learning (ML)? Moreover, to what extent can the tools and methodologies derived from computational linguistics contribute qualitative and quantitative metrics to bolster ongoing empirical research in Assyriology?

The presentation navigates through these overarching questions, with a particular focus on the innovative contributions of the FactGrid Cuneiform Wikibase project ([https://database.factgrid.de/wiki/FactGrid:Cuneiform\\_Project](https://database.factgrid.de/wiki/FactGrid:Cuneiform_Project)). This initiative seeks to unravel the untapped potential of the half-a-million cuneiform sources worldwide, aiming to construct reproducible methods for generating factual (triple) statements about the myriad named entities embedded within these ancient artifacts. Join us on this intellectual journey as we employ quantitative bibliometrics to illuminate the uncharted territories of Assyriology and chart a course towards more nuanced and data-driven scholarship.

### *Digital Philology on the Ancient Silk Road*

Hannes Fellner (University of Vienna)

From the 2nd century CE on, communities and monasteries developed along the trade routes of the ancient Silk Road in and around the Tarim Basin in today's Xinjiang Uyghur Autonomous Region in the People's Republic of China. These were centres of writing, copying, translating, and transmitting texts similar to the monasteries in medieval Europe.

The old Indo-European languages Sanskrit, Tocharian, and Saka written in a Central Asian variant of the Indian Brahmi script – Tarim Brahmi – were the major languages in use in the Tarim Basin in the first millennium CE. In contrast to the writing traditions in medieval Europe, the ones on this part of the Silk Road are not well understood, mainly due to the fragmentary status of texts. In this talk, I will address recent efforts of making these languages and the Tarim Brahmi script digitally accessible and operable for philological, palaeographic, and linguistic research in the framework of the FWF-START project "The characters that shaped the Silk Road – A database and digital palaeography of Tarim Brahmi". In the project, the text witnesses are linked to their digital facsimiles on the character level using Transkribus. All data concerning the texts is combined in an XML database and published through a web application. This allows to determine which text was written by whom, when, where, and how in order to

- trace the evolution of Tarim Brahmi and its adaptation to the different languages
- reveal the relationship between script types, languages, and genres
- categorize countless text fragments that are so far unidentified (regarding language, provenance, date, genre etc.)

- potentially (re)combine scattered fragments belonging to the same manuscript leaf
- and, of course, to better understand literacy and writing culture in the Tarim Basin

### ***Tribute to Silvin Košak***

**Metka Furlan, Luka Brenko (University of Ljubljana)**

This contribution is dedicated to the memory of our fellow compatriot, Silvin Košak, who passed away last year, and whose immense desire to pursue excellence in his field led him from his homeland to the wider world.

His studies in Indo-European Comparative Linguistics and Archaeology at the Faculty of Arts of the University of Ljubljana were influenced by Professors Bojan Čop and Viktor Korošec. As an assistant to Professor Čop, he worked at the Chair of Ancient Oriental Studies of the Department of Comparative Linguistics, established specially for him, from 1969 onwards. During this time, he published the booklet *The King of Battle (Kralj boja, 1971)* and *The Reverse Dictionary of Hittite Names (Odzadni slovar hetitskih imen, 1974)*. After obtaining his PhD in 1975, his desire for better working conditions led him abroad, where he realized his dream and became one of the most renowned Hittite scholars of his time. His life abroad was marked mainly by Chicago, Mainz and Changchun, before settling in Mainz after 1985. The crown of his rich scientific work is undoubtedly the decades-long project of *The Concordance of Hittite Cuneiform Tablets (Konkordanz der hethitischen Keilschrifttafeln)*, which he continued tirelessly even after his retirement in 2007, until his death in early 2022 prevented him from doing so.

In 2009, Silvin Košak was honoured by the Slovenian academia, becoming a Corresponding Member of the Slovenian Academy of Sciences and Arts, while already in 2007, more than 60 colleagues from all over the world expressed their respect and gratitude for his contributions to the field in the Festschrift volume *Tabularia Hethaeorum*.

Although Silvin Košak spent most of his life abroad, he has never forgotten his home country and has generously donated his library to his former department. This extensive corpus of books, articles and other material from the field of Hittitology and ancient Oriental studies is invaluable for the development of the study of Hittite in Slovenia, and as a whole represents not only a comprehensive source of knowledge and a thorough overview of the discipline, but also the man who built up this collection. The many annotations in the books, the enclosed copies, the dedications and the written correspondence will continue to reveal to future generations the mind of a man who touched many with his knowledge and, above all, with his weaving of human bonds.

### ***Ancient Language Processing: A Transformative Framework for the Computational Study of Ancient Texts***

**Shai Gordin (Ariel University, Open University of Israel)**

The field of ancient language processing (ALP) is witnessing an expanding research community, driven by the increasing availability of ancient language resources and the growing interest of scholars with machine learning expertise in this domain. This talk delves into the foundational aspects of ALP by exploring how machines help find patterns in texts through a series of transformations: from image to model, from print to plain text, from words to vectors. The presentation unfolds into specific challenges and their resolution strategies in ALP, leveraging a series of projects ranging from the computational modeling of ancient languages to the intricacies

of deciphering texts with optical character recognition (OCR) and machine translation. Through these case studies I will expose the nuances of AI research in the ancient linguistic landscape.

### ***Classical Languages at the Crossroads of Tradition and Artificial Intelligence***

**Matej Hriberšek (University of Ljubljana)**

In the last fifty years, Classical Philology, in collaboration with other related disciplines, has undergone a remarkable quantitative and qualitative leap forward. This has been due not only to an increased interest in ancient literature, culture and civilization, but also to an increased interest in various aspects of the study of Greek and Latin. The development of computers was the first step towards the creation of the first corpora of texts, followed by a series of software tools that enabled and facilitated more sophisticated research approaches and procedures; these tools are being improved year by year and allow for better research. And what can we expect in the era of the infiltration of AI in the scientific work in the field of classical philology? Its use today is a fact that cannot be avoided, and it is too early to answer the question of how its use will affect research and study. In any case, the complexity of the classical method of language research means that it will not disappear in the near future, since some areas of research require processes of such complexity that technology will not be able to compete with them for a long time. The answer to the challenges posed by artificial intelligence lies in combining classical methods and approaches with the judicious use of new technologies.

### ***The Electronic Babylonian Library: Taming the Cuneiform Jungle***

**Enrique Jiménez (Ludwig-Maximilians-Universität München)**

The Electronic Babylonian Library (eBL) Platform, funded by a Sofja Kovalevskaja prize (Humboldt foundation, 2018–2024), brings together ancient Near Eastern specialists and data scientists to revolutionize the way in which the literature of Iraq in the first millennium BCE is reconstructed and analyzed. Generations of scholars have striven to explore the written culture of this period, in which literature in cuneiform script flourished to an unprecedented degree, but their efforts have been hampered by two factors: the literature's fragmentary state of reconstruction and the lack of an electronic corpus of texts on which to perform computer-aided analyses.

The eBL project aims to overcome both challenges. First, a comprehensive electronic corpus has been compiled, and large amounts of fragments have been transliterated into a database ("Fragmentarium"). A sequence alignment algorithm has been developed to query these corpora. The work will propel the reconstruction of Babylonian literature forward: thousands of new fragments have been identified, over one thousand joins established, and many lacunas in literary texts have been filled.

In a symposium in memory of S. Košak, whose contribution to Hittite bibliography has shaped the field, it seems appropriate to discuss in particular how the eBL platform deals with the very complex history of the publication of cuneiform tabillas, as well as the next steps planned in this regard.

## ***From Filecards to AI. Development and Prospects of Hittitology from a Technical Perspective***

**Gerfrid G.W. Müller (Würzburg University)**

What tools and resources were available to Hittitology when Silvin Kořak moved from the Chicago Hittite Dictionary to Hittite research at the Academy of Sciences and Literature in Mainz? Starting from this perspective, the development steps with all their obstacles and insights that led to the Hittitology Portal Mainz in its present form will be traced. As a living research infrastructure, the portal is constantly changing, which is why new, unfinished projects will also be presented and perspectives for the future addressed. In view of the hype surrounding AI, the question arises as to whether we will still need philologists and linguists in the future. (Or computer scientists? 😊)

## ***Revival of Ancient Sanskrit Teaching Methods using Computational Platforms***

**Preeti Shukla (Central Sanskrit University Delhi), Amba Kulkarni (University of Hyderabad), and Devanand Shukl (Central Sanskrit University Delhi)**

We look at the Ancient Indian methods of learning / teaching viz. Daṇḍānvaya and Khaṇḍānvaya to understand the advantages they have over the Western methods which failed to develop necessary proficiency among students. The ancient methods at the same time demanded good hold over the Pāṇini's grammar which the modern students lack. With the availability of various computational platforms for analysing Sanskrit texts, we then discuss, how we can revive the ancient methods of learning / teaching Sanskrit. And finally we illustrate with an example, how these computational platforms can be harnessed to develop domain specific learning capsules and on-line readers for specific texts. With the available computational platforms it is possible to analyse any Sanskrit text semi-automatically and present the same in a more compact and user-friendly mode. The user interfaces can be tailored to the needs of a student. The on-line readers provide the texts with the following information -

- Saṃhitā-pāṭha (original continuous text)
- Pada-pāṭha (segmented text)
- Pada-viśleṣaṇa (morphological analysis)
- Samasta-pada-viśleṣaṇa (compound word analysis)
- Anvaya (parsed structure), and the possible prose order
- Dictionary help for each word

# Speakers

## Adam Anderson

- B.A. in Linguistics
- Masters in Assyriology at Ludwig-Maximilians University, Munich
- PhD in Assyriology at Harvard University
- Specialization in Old Assyrian Texts with Mogens Larsen at Copenhagen University
- Postdoc Lecturer at UC Berkeley, with a grant from the Mellon Foundation to create a Digital Humanities minor
- Research director through the Data Science Discovery Program at UC Berkeley
- Research director of the FactGrid Cuneiform Project  
([https://database.factgrid.de/wiki/FactGrid:Cuneiform\\_Project](https://database.factgrid.de/wiki/FactGrid:Cuneiform_Project))

## Luka Brenko

Luka Brenko is a Research Assistant at the Department of Comparative and General Linguistics at the Faculty of Arts of the University of Ljubljana. After receiving his Master's degrees in comparative linguistics and French studies in 2020, his main interest has been Hittite and Anatolian in the context of Indo-European. He is pursuing a PhD in historical linguistics with the topic *Hittite Reduplicated Verbal Formations and Their Proto-Indo-European Background*.

## Hannes A. Fellner

Hannes A. Fellner studied linguistics at the University of Vienna and received his PhD from Harvard University in 2013. He was a postdoctoral researcher at the University of Vienna and assistant professor at Leiden University. Since 2018 he is the principal investigator of a START-project funded by Austrian Science Fund dedicated to the research of the Central Asian variants of the Indian Brahmi script. He is currently associate professor for historical linguistics and digital philology at the University of Vienna. He is a member of the Young Academy of the Austrian Academy of Sciences and the director of the Austrian Institute for Research on China and Southeast Asia. His research interests include Indo-European nominal morphology, historical and comparative linguistics and philology of the Indo-European languages of the ancient Silk Road, and theoretical approaches to language change.

## Metka Furlan

Metka Furlan is a Professor of Indo-European comparative linguistics and Scientific Councillor at the Fran Ramovš Institute of Slovenian Language at the Research Centre of the Slovenian Academy of Sciences and Arts. Her work is focused on Slovenian, Slavic and Hittite from the point

of view of Indo-European comparative linguistics, etymology and onomastics. She is co-author of Bezlaj's *Etymological Dictionary of the Slovenian Language*, the author of *New Etymological Dictionary of the Slovenian Language* as a growing online dictionary (2017-), a member of the editorial board of the journal *Jezikoslovni zapiski (Linguistic Notes)*, and a lecturer of the Hittite language and laryngeal theory at the Department of Comparative and General Linguistics at the Faculty of Arts of the University of Ljubljana. Although retired at the beginning of 2023, she continues her work at the faculty and the institute.

## **Shai Gordin**

Shai Gordin is Senior Lecturer for ancient Near Eastern History and Digital Humanities at Ariel University, and a visiting Professor at the Digital Humanities and Social Sciences Hub (DHSS Hub), Open University of Israel. He is the PI of the *Babylonian Engine* project and heads the Digital Past Lab, which conducts research in the computational analysis of ancient texts and artefacts (<https://digitalpasts.github.io/docs/projects.html>). He is the co-initiator of the digital ancient Near Eastern studies network (DANES), and co-organizer of the Ancient Language Processing workshop at RANLP2023, and the first international shared task of ancient cuneiform languages translation in MT-SUMMIT2023. He currently serves on the editorial board of PLOS ONE.

## **Matej Hriberšek**

Assoc. Prof. Dr. Matej Hriberšek (born 1972 in Celje, Slovenia) studied Classical Philology at the Faculty of Arts, University of Ljubljana (1991-1996; MA 2000; PhD 2003). He became an assistant at the Department of Classics in 1997, and an assistant professor in 2010. His main fields of interest are Latin and Greek grammar and metrics, history of philology, history of linguistics, history of education, medieval and Neo-Latin literature, lexicography (main editor of Latin-Slovenian and Greek-Slovenian dictionary), and translating from classical languages. In 2008-2010 he was president of the Slovenian Association of Literary Translators, in 2010-2012 president of the Administrative Board of the Faculty of Philosophy in Ljubljana, and in 2017 visiting professor at the Institute of Slavonic Studies at the Karl-Franzens University in Graz. He leads the research project "Latin and German Chronicles in Slovenia" and from 2021 he serves as Vice Dean of the Faculty of Arts in Ljubljana.

## **Enrique Jiménez**

Prof. Dr. Enrique Jiménez holds a chair in Ancient Near Eastern Literatures at LMU since 2018 and is the winner of a 2017 Sofja Kovalevskaja Award. He specializes in the literature and scholarly texts from ancient Mesopotamia, in particular from the first millennium BCE. He is co-founder and co-director of the "Cuneiform Commentaries Project" (<https://ccp.yale.edu/>), funded by Yale University (2013–2016) and the National Endowment for the Humanities (Division of Research

Programs “Scholarly Editions and Translations,” 2015–2018). He is PI of the “Electronic Babylonian Literature” project, and founder of the “electronic Babylonian Library” platform (<http://www.ebl.lmu.de/>), a tool that has propelled the reconstruction of Babylonian literature. He is co-PI of the project “Cuneiform Artefacts of Iraq in Context” (CAIC) at the Bayerische Akademie der Wissenschaften (BAW, 2022–2046), which is based on the eBL platform.

## **Amba Kulkarni**

Amba Kulkarni is a computational linguist. Since 1991 she has been engaged in showing the relevance of Indian Grammatical Tradition to the field of computational linguistics. She has contributed towards the building of Anusaarakas (language accessors) among English and Indian languages. She is the founder head of the Department of Sanskrit Studies, University of Hyderabad established in 2006. Since then her focus of research is on use of Indian grammatical theories for computational processing of Sanskrit texts. Under her leadership, a Consortium of institutes developed several computational tools for Sanskrit and also a prototype of Sanskrit-Hindi Machine Translation system. In 2015, she was awarded a ‘Viśiṣṭa Sanskrit Sevāvratī sammāna’ by Central Sanskrit University (erstwhile Rashtriya Sanskrit Sansthan), New Delhi for her contribution to the studies and research on Sanskrit based knowledge system. She was a fellow at Indian Institute of Advanced Study, Shimla from 2015-17. She has 9 books authored/edited to her credit, with 73 journal and conference papers. She also executed several projects on Computational Linguistics receiving funding from India and abroad. She has guided 9 doctoral students.

## **Gerfrid G.W. Müller**

Gerfrid G.W. Müller studied Assyriology, Semitic Studies and Economics in Heidelberg and Cuneiform Studies, Semitic Studies, Japanese and Economic History in Hamburg. He received his Ph.D. from the University of Heidelberg and habilitated in Ancient Near Eastern Studies at the University of Münster, specializing in Akkadian Studies, Economic and Social History of the Ancient Near East, Historical Geography, and Digital Infrastructures for Cuneiform Texts. He has also been a senior research associate and adjunct professor at the Würzburg University, held a term professorship at the Free University of Berlin, taught at the University of Cairo and at Suor Orsola Benincasa in Naples, co-founded the Heidelberg Orientverlag, and worked as project manager, project coordinator, or project team member on several major Hittite projects, including the *Hethitische Bibliographie*, *Ortsnamen der Hethiter*, *3D-Joins und Schriftmetrologie*, and *Corpus der Hethitischen Festrutuale*. Currently his work focuses on *Computer-unterstützte Keilschriftanalyse* and *Thesaurus Linguarum Hethaeorum digitalis (TLHdig)*. Since 2001 he has been the technical director and organizer of the Hethitologieportal Mainz, and since 2007 he has been an expert in the IT Commission of the Academy of Sciences and Literature in Mainz.

# **Devanand Shukla**

## **Job profile**

Deputy Director (Academic)

## **Area of Specialization**

Sanskrit (Paninian Vyakarana)

Sanskrit Studies (Sanskrit Science and Technology)

Indian Knowledge System and Sanskrit Knowledge Systems

Natural Language Processing (with Specialization in Language Technology)

Sanskrit Studies (Sanskrit Computational Linguistics)

Indian Lexicography

E-Learning and E-Corpora

E-Content Development

## **Computer Proficiency**

Linux and Windows operating systems.

Scripting Languages: XML, HTML, Javascript.

Document Processing Systems: LaTeX, Indesign.

Machine Translation Tools especially Sanskrit Computational Tools.

## **Educational Qualification**

Ph.D. in Sanskrit Grammar and Manuscriptology

NET-JRF

Post Graduation in Sanskrit Grammar and Sanskrit Literature

## **Employment Scan**

Total Working Experience – 17 Years

Working as Deputy Director (Academic) from 2017 till now in Central Sanskrit University, New Delhi

Worked as Programme Officer (Academic) from 2011 to 2017 in Maharshi Sandipani Rastriya Vedavidya Pratisthan (Under the Ministry of HRD Govt of India), Ujjain

Worked as a Senior Linguist from 2008 to 2011 in a project entitled 'Development of Computational Toolkit and Sanskrit-Hindi Machine Translation System', in Department of Sanskrit Studies, University of Hyderabad, Hyderabad

Worked as a Acharya from April 2007 to May 2008 in Gurukulam Vimukti Vidyapeeth, Mumbai

Worked as a Lecturer in Sanskrit from August 2006 to March 2007 in Post Graduate Studies & Research, S.N.D.T. Women's University, Mumbai

Worked as a Lecturer in Vyakarana from Dec. 2005 to August 2006 in K.J.Somaiya Sanskrit Vidyapeetham, Rastriya Sanskrit Sansthan, (Deemed University), Mumbai

## **Books Published 16**

### **Paper Published**

International Journals 03

National Journals 17

Papers Presented in Conference/Symposium/Seminars/Workshops

International 06

National 06

Workshops Participated as a Resource Person 12

## **Preeti Shukla**

### **Job Profile -**

Currently working –

· Assistant Professor (Guest) in Institute of Open and Distance Education (Central Sanskrit University)

Previously worked -

· Language Editor in the Project 'Machine Translation System' sponsored by Language Technology and Research Center, International Institute of Information Technology, Hyderabad.

· Project Fellow in Samskrit Promotion Foundation, Delhi.

### **Educational Qualifications -**

· Ph.D. in Sanskrit Computational Linguistics

· M. Phil in Shabdabodha Systems and Language Technologies

· NET-JRF

· P G Diploma in Comparative Mythology

· Post Graduation in Sanskrit and English

### **Research Interests -**

· Indian Philosophy (Darshana) and Language Philosophy (Shabdabodha Systems)

· Indian Knowledge Systems and Sanskrit Knowledge Systems

· Sanskrit Studies (Sanskrit Computational Linguistics)

· Natural Language Processing (Language Technology)

· Pedagogy (Indian and Western)

· Indian Lexicography

· E-Learning and E-Corpora

· E-Content Development

### **Computer Proficiency -**

· Linux and Windows operating systems.

· Scripting Languages: XML, HTML.

· Document Processing Systems: LaTeX.

· Machine Translation Tools especially Sanskrit Computational Tools.

### **Books Authored -**

1. Shreemadbhagavadgeeta Shankari-vivriti

2. Shreemadbhagavadgeeta Padartha Kosh (*Awarded "Shankar Puraskar" by Uttar Pradesh Sanskrit Sansthanam*)

3. Kāraṅkānirūpaṇam

**Several research papers published in National and International Journals.**

### **Participated as a Resource Person in -**

· Various Sanskrit Computational Linguistics workshops.

· NCERT workshops on 'Bhasha Sangam', 'Ek Bharat Shreshtha Bharat', 'Samajh ka Madhyam', etc.