International Journal of Social Science Research (IJSSR)
Volume- 3, Issue- 1 | January - February 2026 ISSN: 3048-9490

Samsaadhanii: A Comprehensive Platform for Sanskrit Computational Linguistics

Vishal Sharma[1] & Bhushan Gautam[2]
[1]Research Scholar, Central Sanskrit University, Ved Vyas Campus, Balahar
[2]Research Scholar, Central University of Himachal Pradesh, Dharamshala

Abstract

Samsaadhanii is an innovative computational platform designed to process and analyze the Sanskrit language. Developed at the University of Hyderabad under the guidance of Professor Amba Kulkarni, this system bridges the gap between ancient linguistic traditions and modern computer science. Sanskrit, often called the language of the gods, possesses a highly complex structure characterized by rich word formations (morphology), strict sound blending rules (*Sandhi*), and flexible word order. These unique features make it difficult for standard computer programs and translation tools to interpret the language correctly. This paper provides a detailed overview of *Samsaadhanii*, describing its history, technical architecture, and practical applications. Unlike many modern AI tools that guess meanings based on statistics, *Samsaadhanii* uses a rule-based approach. It digitizes the precise grammatical rules formulated by the ancient scholar Panini, effectively treating hisgrammar as computer code. The paper highlights the platform's key tools, including the Morphological Analyzer, which breaks words down into their roots, the *Sandhi* Splitter, which separates merged words, and the Parser, which analyzes sentence structure using traditional *Karaka* theory. The study concludes by discussing the platform's significant impact on education and research. By providing free, open-source access to these powerful tools, *Samsaadhanii* plays a crucial role in preserving cultural heritage and making the vast knowledge contained in ancient Sanskrit texts accessible to the modern digital world.

Keywords: Sanskrit computational linguistics, *Paninian* grammar, Morphological analysis, Sandhi tools, Digital humanities, Language processing.

Introduction

Sanskrit is often referred to as *Devavani* (the language of the gods) and is one of the oldest recorded languages in human history, with a literary tradition spanning over 3,500 years. It serves as the primary vessel for a vast amount of knowledge, including Indian philosophy, mathematics, astronomy, medicine (*Ayurveda*), and literature. However, in the modern digital age, accessing this knowledge requires more than just reading books; it requires computational processing, the ability for computers to read, search, translate, and analyze these texts automatically.

*Corresponding Author Email: vishalsharma21j@gmail.com
Published: 14 February 2026
DOI: https://doi.org/10.70558/IJSSR.2026.v3.i1.30836

The challenge is that Sanskrit works very differently from languages like English. English relies heavily on word order (Subject-Verb-Object) to convey meaning. Sanskrit, however, is a highly inflectional language. This means that the endings of words change to show their meaning, and the words can be placed in almost any order in a sentence without changing the meaning. Additionally, Sanskrit uses a system called *Sandhi*, where the sounds of two words merge together when they are spoken or written next to each other. Because of these complexities, standard tools like Google Translate or basic spell-checkers often fail when applied to Sanskrit. To solve this, the field of Sanskrit Computational Linguistics was born. Among the various tools developed, *Samsaadhanii* stands out as a premier platform. Developed at the University of Hyderabad, it uses the precise logic of ancient grammar rules combined with modern software engineering to teach computers how to understand Sanskrit.

History and Development

The initiative began around 2002. At this time, Professor Amba Kulkarni, a pioneer in the field who holds backgrounds in both Computer and Sanskrit, recognized that statistical methods (guessing meanings based on probability) would not work well for Sanskrit. Instead, she proposed using the rigorous rules laid out by *Panini*[1]. Early work focused on basic machine translation and building a morphological analyzer, a tool that identifies what a word is. In 2006, the Department of Sanskrit Studies was established at the University of Hyderabad, which became the home for this project. The university provided the necessary academic environment and funding grants to expand the project from simple experiments to a robust system. During this phase, students from Rashtriya Sanskrit Vidyapeetham in Tirupati contributed significantly, particularly in building tools for *Sandhi* (splitting and joining words). Over the last two decades, *Samsaadhanii* has evolved into a fully integrated web platform. It is no longer just a research experiment but a public utility available to the world. As of 2025, the platform is open-source (available on GitHub), allowing programmers globally to contribute. It has hosted major international conferences and is regularly updated to include new texts and improved algorithms.

Theoretical Basis

What makes *Samsaadhanii* unique is that it thinks like a Sanskrit scholar, not like a modern chatbot. It is built upon three major theoretical pillars derived from Indian tradition. The core operating system of *Samsaadhanii* is the *Ashtadhyayi*, written by the grammarian Panini around the 4th century BCE. Panini created about 4,000 rules (*sutras*) that function like computer code. These rules describe exactly how to build a valid Sanskrit word from a root sound. *Samsaadhanii* digitizes these rules. If the computer wants to know if a word is correct, it checks these 4,000 rules.

In English grammar, we look for the Subject and the Predicate. in Sanskrit, the focus is on *Karaka*, which describes the relationship between a noun and the action (verb).

[1] Panini- An ancient Sanskrit philologist, grammarian, and revered scholar in ancient India, traditionally dated between the 6th and 4th centuries BCE. He is often considered the "father of linguistics" and arguably the first computer scientist because his grammar rules function like an algorithm.

- ***Karta*:** The doer of the action.
- ***Karma*:** The object of the action.
- ***Adhikarana*:** The location of the action.

Since Sanskrit word order is free, *Samsaadhanii* uses *Karaka* theory to draw lines of connection between words, regardless of where they sit in the sentence. This is called Dependency Parsing. *Sabdabodha* is the theory of how a human mind understands a sentence. It involves looking at the words, remembering their meanings, and understanding their connection. *Samsaadhanii* attempts to replicate this cognitive process step-by-step to derive the final meaning of a text.

Key Features and Tools

Samsaadhanii is built with a modular design, meaning users can use individual tools or combine them. It is web-based, with mobile apps for some features, and can be installed locally via Docker.

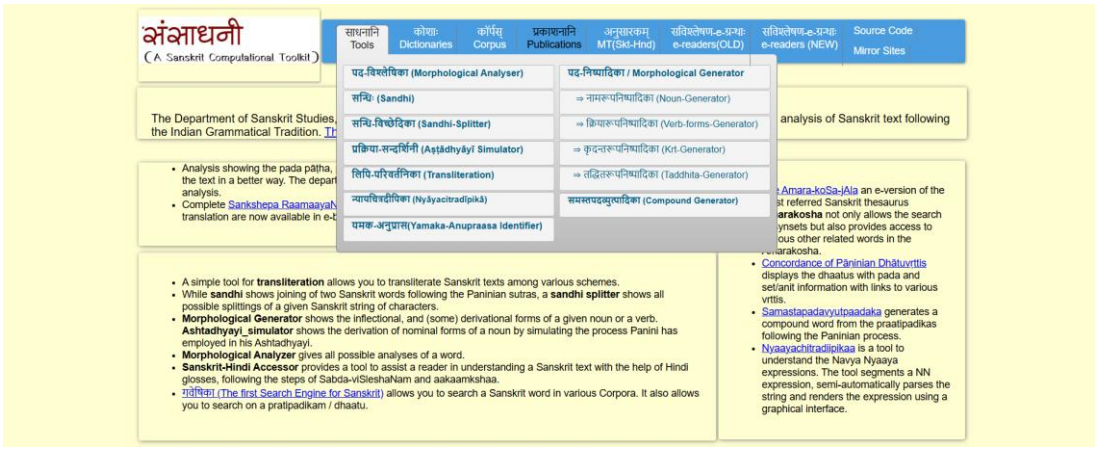
Key Tools:

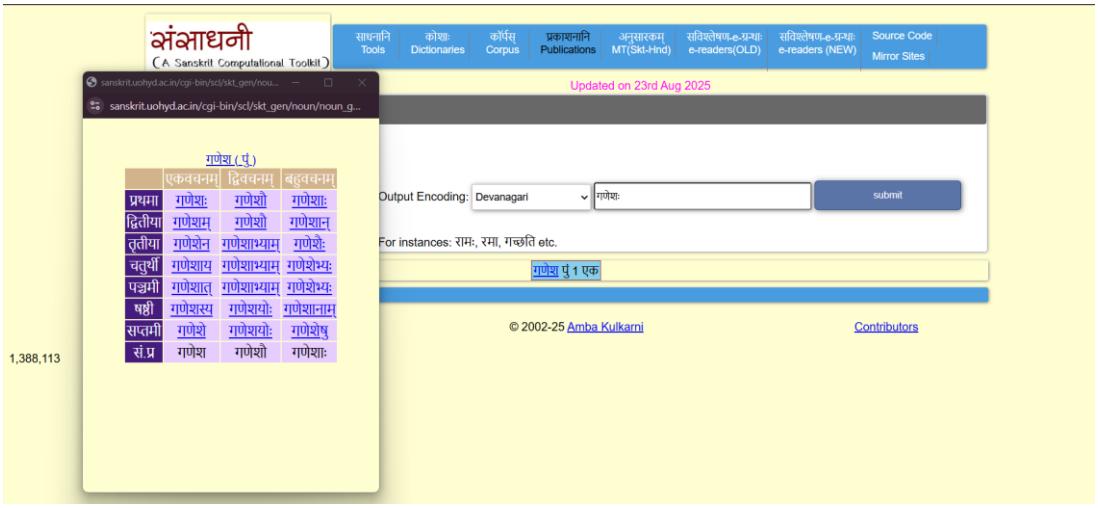
- **Morphological Analyzer and Generator**: Breaks down Sanskrit words into their root, grammar details (like tense, number, case), and creates different forms of a word. This handles thousands of possible changes in Sanskrit words.
    - The analyzer takes a Sanskrit word like *Ramah* and lists all possible roots, endings, and meanings.
    - The generator does the opposite, given a root and grammatical details, it creates the correct word form. There's also a verb generator and noun declension tool.
- **Sandhi[2] Splitter and Joiner**: Sandhi is when sounds change or merge at word boundaries. The tool splits joined words or joins them, following *Panini's* rules, and handles ambiguities by suggesting options.
    - In Sanskrit, words blend when *Sandhi*, like *deva* + *ishah* becomes *deveshah*.
    - The splitter finds all possible ways to separate a combined string.
    - The joiner applies rules to combine words correctly.
- **Parser:** Analyzes sentence structure using dependency grammar, focusing on relationships between words rather than fixed order. It uses concepts like *karaka* from traditional grammar.
- **Transliteration Tools:** Converts text between *Devanagari* script and Roman schemes like IAST or Harvard-Kyoto. A mobile app is available for quick transliteration.

Machine Translation: Includes Anusaaraka system for translating Sanskrit to Hindi, using rule-based methods instead of pure statistics. It provides a tool to assist a reader in

[2] A linguistic term for the process where sounds change or merge at the boundaries of words .In Sanskrit, these rules are mandatory and very strict.

understanding a Sanskrit text with the help of Hindi glosses, following the steps of *Sabda-viSleshaNam* and *Aakaamkshaa*. Other features include segmenters for compounds and tools for prose order normalization.





Other Useful Tools

- Amarakosha search is an electronic version of the famous Sanskrit thesaurus for synonyms.
- Gaveshika is a search engine for Sanskrit words in corpora large text collections.
- Compound word generator *Samasa*.
- Tools for *Navya Nyaya* with advanced logic expressions.

Architecture and Technical Details

Samsaadhanii relies heavily on a computer science concept called Finite State Automata or Transducers. Imagine a flowchart. The computer starts at State A. If it sees the letter a,it moves to State B. If it sees b, it moves to State C. By building massive flowcharts that represent *Panini's* rules, the system can process words very quickly and with 100% grammatical accuracy, provided the rules are coded correctly.

The system is modular. This means the *Sandhi* Splitter is a separate block from the Dictionary.

If the team wants to update the dictionary, they don't have to break the *Sandhi* tool. This makes maintenance easier.

The logic needs data to work on. Samsaadhanii utilizes:

- **Dhatupatha:** A database of roughly 2,000 verb roots.
- **Amarakosha:** Lexical database for nouns.
- **Annotated Corpora:** Collections of texts that have been manually checked by humans to serve as a "gold standard" for testing.

This architecture ensures high accuracy for classical Sanskrit, where rules are strict, rather than relying on probability like in modern languages.In simple terms, the computer follows the same logical steps a human Sanskrit scholar would, but much faster.

Applications

Samsaadhanii is not just for computer scientists; it has real-world uses across various fields. Scholars studying ancient texts often have to search through thousands of pages to find specific patterns. *Samsaadhanii* allows them to query a text. For example, a researcher could ask, Show me every time the word Dharma is used as an object in the Mahabharata. This type of analysis was impossible before digital tools.

For students learning Sanskrit, the platform acts as a 24/7 tutor. If a student is stuck on a complex word in a textbook, they can type it into the analyzer. The tool will peel back the layers of grammar, explaining the root and the suffix. This builds confidence and speeds up the learning process.

Many Sanskrit manuscripts are deteriorating. By digitizing them and using *Samsaadhanii* to make them searchable, we preserve the contents for eternity. The platform aids in creating E-Readers where a user can click on any word in a digital book to see its meaning.

While it cannot yet replace a human translator for poetry, it serves as a powerful assistant. It provides the skeleton of the meaning, allowing the human translator to flesh out the artistic nuance without getting bogged down in complex grammar.

Accessibility

The platform is user-friendly:

- Web interface at scl.samsaadhanii.in or related sites.
- Mobile apps for transliteration.
- Docker for local installation.
- GitHub for source code and contributions.

It is free and open, encouraging community input.

Impact and Recognition

Samsaadhanii has transformed Sanskrit studies. It has led to many publications, conferences like World Sanskrit Conference sections on computational linguistics, and international

collaborations. In education, it makes Sanskrit more approachable, sparking interest among young learners. For cultural preservation, it aids global digitization efforts.

Other initiatives exist, like Gérard Huet's Sanskrit Heritage site in France or projects in Germany and the US. *Samsaadhanii* stands out for its strong use of *Paninian* rules and focus on traditional theories, unlike more statistical approaches elsewhere.It influences tools for other ancient languages.

Challenges and Limitations

Despite its strengths, Samsaadhanii faces challenges:

Sanskrit is a language of double meanings. A single string of letters can sometimes be split in three different ways, all of which are grammatically correct but only one of which makes sense in the story.

- *Example:* The string *svetodhavati* could mean "The white dog runs" or "The steam runs." Currently, the tool lists all options. It requires a human to choose the right one. This is called the problem of disambiguation.

The tools are optimized for Classical Sanskrit (the standardized language of *Panini*). However, older *Vedic Sanskrit* or later regional poetic variations may not follow these rules strictly. The tools sometimes fail to analyze these irregular texts. While powerful, the interface is designed by academics for academics. It can be intimidating for a casual user or a beginner. Simplifying the visual design and making it more "app-like" is a goal for the future. Processing huge texts (like the *Ramayana*) requires significant computing power. The algorithms need to be optimized to run faster on standard computers.

Conclusion

Samsaadhanii is a remarkable bridge between ancient Sanskrit wisdom and modern technology. Led by Professor Amba Kulkarni for over two decades, it provides essential tools for analyzing this complex language. It supports research, education, and preservation, making Sanskrit more accessible in the digital age. From helping beginners read classics to aiding advanced research, it has proven the practicality of Panini's grammar in digital form. With ongoing developments, collaborations, and growing user base (including apps), it promises to play a key role in Sanskrit revival. Platforms like this show that computational linguistics can honor cultural heritage while pushing technological boundaries. As more people discover *Samsaadhanii,* it could inspire similar efforts for other classical languages worldwide. By staying true to traditional grammar while embracing computation, *Samsaadhanii* shows how technology can enrich humanities. Its ongoing development promises even greater impact, inspiring similar efforts for other classical languages. In a world of fast-changing tech, platforms like this ensure ancient knowledge remains alive and relevant.

References.

1. *Kulkarni, A., & Shukl, D. (2009). Sanskrit Morphological Analyzer: Some Issues and Solutions. Bhartiya Anuvad Parishad.*

2. *Huet, G., Kulkarni, A., & Scharf, P. (Eds.). (2009). Sanskrit Computational Linguistics: First and Second International Symposia. Springer.*
3. *Kulkarni, A. (2019). Sanskrit parsing: Based on the theories of śābdabodha. D.K. Printworld*
4. *Joshi, S. D., & Roodbergen, J. A. F. (1991). Pāṇini's grammar and modern linguistics. Sahitya Akademi.*
5. *University of Hyderabad, Department of Sanskrit Studies. Samsaadhanii: Sanskrit computational linguistics tools. https://sanskrit.uohyd.ac.in/scl/*
6. *Pāṇini. (2002). Aṣṭādhyāyī (S. D. Joshi & J. A. F. Roodbergen, Eds.). Sahitya Akademi.*