DAILY EDITION · 25 NOVEMBER 2026

Sanskrit & AI News · 25 November 2026

25 November 2026 · Daily headline digest from reliable sources · 9 headlines in 4 of 8 categories, in order 01–08 · headline, link, source and date, with a one-line brief · Curated by Dr P.R.K. Varma

Published 25 November 2026, 08:30 IST · Report a correction

Items first reported between 18 and 24 November 2026, each opened and checked on 25 November 2026. · About 3 min read · PDF · RSS

DAILY SANSKRIT & AI NEWS

Sanskrit AI Universe

25 November 2026

9 headlines · 4 of 8 categories active

07 Manuscripts & Heritage 3 · 06 Benchmarks & Data 2 · 03 Knowledge & Meaning 2

01 Grammar & Computation · 02 Sanskrit NLP & LLMs · 03 Knowledge & Meaning · 04 Metre & Sound · 05 Mind & Philosophy · 06 Benchmarks & Data · 07 Manuscripts & Heritage · 08 Events & Calls

Curated by Dr P.R.K. Varma

Today in 60 seconds

1. MIDF/eGangotri Sanskrit manuscripts dataset on Hugging Face adds machine-generated OCR for 182,417 pages (Hugging Face (ArthatvBhujade))
   06 · Benchmarks & Data
2. Sanskrit Wikisource corpus of 2,653 works released in IAST plain text (Zenodo)
   06 · Benchmarks & Data

3. BHU's Saraswatam centre sets out 243 Sanskrit manuscripts for conservation and digitisation (Dainik Bhaskar) 07 · Manuscripts & Heritage
4. Preprint reports a failed attempt to date the Bṛhadāraṇyaka Upaniṣad with a 'linguistic clock' (Zenodo (Atharva Chourey Lab)) 03 · Knowledge & Meaning
5. Small ByT5 model for correcting Sanskrit OCR errors released on Hugging Face (Hugging Face (MrunalYewatkar15)) 02 · Sanskrit NLP & LLMs

DISCOVERY OF THE DAY

MIDF/eGangotri Sanskrit manuscripts dataset on Hugging Face adds machine-generated OCR for 182,417 pages

The dataset's Smriti OCR v1 holds machine-generated transcriptions of 182,417 pages across 1,488 canonical Devanāgarī manuscripts, with page text, source references, model provenance and review flags; the licence is listed as 'other', so check it before reuse.

Hugging Face (AtharvBhujade) · 18 Nov 2026 · Dataset

0 01 Grammar & Computation	2 02 Sanskrit NLP & LLMs	2 03 Knowledge & Meaning	0 04 Metre & Sound
0 05 Mind & Philosophy	2 06 Benchmarks & Data	3 07 Manuscripts & Heritage	0 08 Events & Calls

PULSE IN 3D

SŪTRA OF THE DAY

अदेङ् गुणः

adeṅ guṇaḥ · Aṣṭādhyāyī 1.1.2

The vowels a, e and o are called guṇa.

The second sūtra of the Aṣṭādhyāyī follows yesterday's vṛddhi (1.1.1) and defines the other vowel grade used throughout its rules; adeṅ is a compound of a and eṅ, the latter standing for e and o.

Read 1.1.2 on ashtadhyayi.com ↗

Editor's note: Second daily edition. Genuine Sanskrit-and-AI news was again thin: nine verified headlines, mostly datasets, tools and manuscript-conservation reports, and none in Grammar & Computation, Metre & Sound or Events & Calls. Several candidates were dropped because they were not new: the Ministry of Culture's Gyan Bharatam conference pages turned out to date from September 2025, and the ISCLS 2027 pages carry deadlines that have already passed. Two features from the journal's own verified notes are labelled as archive pieces, not news.

▶ Evidence labels explained

01 Grammar & Computation

No verified item today.

02 Sanskrit NLP & LLMs

Small ByT5 model for correcting Sanskrit OCR errors released on Hugging Face

Hugging Face (MrunalYewatkar45) · 21 Nov 2026 · Software · New today

Fine-tuned from google/byt5-small on 67,984 noisy–clean line pairs, it lowers character error on five real Tesseract page images from 9.6% to 8.2%; the model card itself warns that gains on real OCR are modest and that about 11% of correct lines get changed.

'From Shloka to Shabda': Sanskrit-to-Bengali machine translation with a new 31,639-pair corpus

ACM Transactions on Asian and Low-Resource Language Information Processing · 19 Nov 2026 · Peer-reviewed · New today

Published in ACM Transactions on Asian and Low-Resource Language Information Processing, the paper fine-tunes mBART and reports BLEU 21.11 and chrF 41.41 on its test set.

03 Knowledge & Meaning

Preprint reports a failed attempt to date the Bṛhadāraṇyaka Upaniṣad with a 'linguistic clock'

Zenodo (Atharva Chourey Lab) · 21 Nov 2026 · Preprint · New today

Before dating the text, the pre-registered pipeline had to pass a blind benchmark; it recovered 0 of 17 control dates (bar: 15) and scored Kendall's τ = 0.000 on ordering the Vedic strata (bar: 0.80), so no date was produced. Not yet peer-reviewed.

Indian Air Force and Central Sanskrit University sign MoU on Yoga learning and certification for Agniveervayu

Press Information Bureau (Ministry of Defence) · 19 Nov 2026 · Official release · New today

Signed at Air Headquarters, the MoU gives Agniveervayu access to structured learning and certification in Yoga through the University.

04 Metre & Sound

No verified item today.

05 Mind & Philosophy

No verified item today.

06 Benchmarks & Data

MIDF/eGangotri Sanskrit manuscripts dataset on Hugging Face adds machine-generated OCR for 182,417 pages

Hugging Face (AtharvBhujade) · 18 Nov 2026 · Dataset · New today

The dataset's Smriti OCR v1 holds machine-generated transcriptions of 182,417 pages across 1,488 canonical Devanāgarī manuscripts, with page text, source references, model provenance and review flags; the licence is listed as 'other', so check it before reuse.

Sanskrit Wikisource corpus of 2,653 works released in IAST plain text

Zenodo · 23 Nov 2026 · Dataset · New today

Roberto de Andrade Martins's Zenodo dataset converts the 16 November 2026 Sanskrit Wikisource dump into about 39.8 million words of IAST text, one file per work, under CC BY-SA 4.0; the scripts were written with an AI assistant.

07 Manuscripts & Heritage

BHU's Saraswatam centre sets out 243 Sanskrit manuscripts for conservation and digitisation

Dainik Bhaskar · 23 Nov 2026 · News report · New today

Dainik Bhaskar reports that the Faculty of Sanskrit Vidya and Dharma Vijnan holds 243 manuscripts (72 on Nyāya, 28 on Vedic literature, 23 on grammar), including a 558-year-old krama-pāṭha of the Vājasaneyi Saṃhitā; digitisation is planned.

Kurukshetra University team surveys 250 rare manuscripts at a Bhiwani Sanskrit college, with a manuscript sub-centre planned

Amar Ujala · 24 Nov 2026 · News report · New today

Amar Ujala reports that the team visited the Haryana Shekhawati Sanskrit Mahavidyalaya library on Monday 23 November; the manuscripts are to be preserved and scanned once a sub-centre is set up.

Karan Singh reviews manuscript conservation at the Ranbir Singh Sanskrit Research Institute in Jammu

Kashmir News Service · 20 Nov 2026 · News report · New today

On 20 November he visited the Institute's Rare Manuscripts Repository at the Raghunath Temple complex, linked by KNS to the Gyan Bharatam initiative, and stressed preserving India's knowledge traditions.

08 Events & Calls

No verified item today.

✦ From the archive · Features

Not news. Pieces drawn from the journal's own verified notes, labelled as such.

FROM THE ARCHIVE · FEATURE

Hidden Grammar in the Stories

Patañjali quoting Kṛṣṇa theatre, a demoness as the grammarians' test case, a charioteer who out-argues a grammarian: 29 verified story–sūtra links.

FROM THE ARCHIVE · SOFTWARE NOTE

Watching a Word Being Born

Twenty rule applications take bhū to bhavati, traced with the Vidyut engine. Step through them one sūtra at a time.

TODAY'S WATCH



Rigveda : Pratishakhya Pratham Patal - 1

CH 01: CEC: SANSKAR: Language and Literature...

2K views · 10 Jan 2021

Why it is here: Not named in the video description (SWAYAM PRABHA UG Sanskrit course, BA 2nd year). UGC/CEC SWAYAM PRABHA channel for UG Sanskrit: a degree-course lecture on the first paṭala of the Ṛgveda Prātiśākhya, the Vedic phonetics manual.

Lead: Small audience and the lecturer is not named in the description; treat as a lead for the course content, not a scholarly reference.

Pairs with today's guṇa sūtra: a university lecture on the Ṛgveda Prātiśākhya, the old Vedic phonetic manual.

All 399 videos in Watch →

? Daily quiz

Answers come only from today's verified items.

1. How many pages does the Smriti OCR v1 release in the MIDF/eGangotri Sanskrit manuscripts dataset cover?

   ▶ Show the answer

2. What did the Bṛhadāraṇyaka Upaniṣad dating preprint produce after failing its blind benchmark?

▶ Show the answer

3. How many works are in the new Sanskrit Wikisource corpus in IAST?

▶ Show the answer

News summaries for information only, not scholarly, religious or professional advice. Each headline links to its original source; read the source before citing it. Items labelled Preprint, Dataset or Software are not peer reviewed.

॥ इति ॥

iti · here ends today's edition.

वागर्थाविव सम्पृक्तौ · *vāgarthāv iva sampṛktau* · Raghuvaṃśa 1.1

← 24 November 2026 · All editions

Sanskrit AI Universe · *Where the Aṣṭādhyāyī meets the algorithm* · वागर्थाविव सम्पृक्तौ

Founding Editor: Dr P.R.K. Varma, MD, DM · Bhimavaram, Andhra Pradesh · ORCID 0009-0002-4674-4403

Sanskrit AI Universe publishes a daily Sanskrit & AI news edition: every headline is checked and linked to its source. The peer-reviewed journal is planned: no ISSN or DOI has yet been assigned, it is not indexed in scholarly databases, no editorial board has yet been appointed, and features are not peer reviewed.

Journal content by Dr P.R.K. Varma is in the public domain (CC0 1.0). Anyone may copy, share, adapt or use it for any purpose, without asking. Linked third-party works keep their own licences.