IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 6, Issue 10, June 2026

Impact Factor: 8.2

The Role of AI in the Indian Knowledge System for Multi-Domain Transformation and Innovation

Anupma Malhotra[1], Dr. Sonam Bansal[2], Dr. Mukti Sahu[3], Parmjeet Kaur[4]

Assistant Professor, Department of Computer Science, Dr. Bhim Rao Ambedkar University, Agra, India[1]

Associate Professor, Department of Education, Gurugram University, Gurugram, India[2]

Lecturer, Department of Political Science

Boudh Panchayat College, Boudh (Affiliated to Sambalpur University), Odisha, India[3]

Assistant Professor, Department of English and Foreign Languages, GSSDGS Khalsa College, Patiala, India[4]

anupmasmalhotra@gmail.com, sonambansal@gurugramuniversity.ac.in

muktisahupolsc@gmail.com, parmjeeteng04@gmail.com

Abstract: *The Indian Knowledge System (IKS) is a vast, time-tested body of codified and tacit knowledge spanning health, agriculture, language, education, the arts, architecture, governance, astronomy and mathematics. Although these traditions encode sophisticated empirical and theoretical reasoning, much of the corpus remains undigitised, linguistically inaccessible and disconnected from contemporary computational pipelines. This paper examines how Artificial Intelligence (AI) can act as an enabling layer for the preservation, interpretation, validation and large-scale dissemination of IKS, and thereby drive transformation and innovation across multiple domains. Adopting a mixed-methods design, we assembled a corpus of 412 records and, after structured screening, analysed 168 eligible studies through parallel bibliometric and thematic tracks. The synthesis is organised around a proposed AI–IKS integration framework that maps eight knowledge domains against six families of AI capability—machine learning, natural language processing and large language models, computer vision, knowledge graphs, generative AI and reasoning systems. Results indicate sustained growth in scholarly attention, an uneven distribution of effort that concentrates in health, agriculture and language technologies, and a persistent gap between current adoption and assessed transformation potential in every domain studied. We further surface cross-cutting enablers and barriers—corpus digitisation, low-resource language modelling, epistemic fidelity, provenance and benefit sharing—and translate them into a research and policy agenda. The study contributes a consolidated multi-domain map, a readiness assessment, and a set of testable propositions intended to guide responsible, culturally grounded deployment of AI within the Indian Knowledge System.*

Keywords: Indian Knowledge System; Artificial Intelligence; natural language processing; Sanskrit computational linguistics; Ayurveda informatics; digital heritage; knowledge graphs; responsible AI; multi-domain innovation

I. INTRODUCTION

The **Indian Knowledge System (IKS)** denotes the accumulated, systematically organised knowledge that has been generated, transmitted and refined across the Indian subcontinent over several millennia. It is not a single discipline but an interconnected ecology of traditions: the medical sciences of Ayurveda, Siddha and Yoga; agronomic and ecological practices encoded in regional almanacs; the grammatical and phonological precision of Sanskrit and other classical languages; the performing and visual arts; temple and civic architecture governed by Vastu and Sthapatya Veda; treatises on polity and ethics; and the mathematical and astronomical contributions associated with figures such as Aryabhata, Brahmagupta and the Kerala school. Much of this material is simultaneously theoretical and empirical, embedding observation, classification and rule-based reasoning that anticipate ideas now central to modern computation.

Two developments have brought renewed urgency to the study of IKS. First, the National Education Policy 2020 explicitly calls for the integration of Indian knowledge traditions into mainstream curricula, and dedicated institutional structures have been created to coordinate research, digitisation and pedagogy around these traditions. Second, the maturation of **Artificial Intelligence**—particularly large language models, computer vision and knowledge-graph technologies—has lowered the cost of working with unstructured, multilingual and heterogeneous knowledge at scale. The conjunction of a policy mandate and a capable technical substrate creates a timely opportunity to ask how AI can serve IKS, rather than merely extract from it.

Yet the opportunity is accompanied by genuine difficulty. A large fraction of the IKS corpus survives as palm-leaf and paper manuscripts that are undigitised, fragile and written in scripts and registers for which annotated data is scarce. Knowledge is frequently tacit, transmitted orally through guru–shishya lineages and embedded in practice rather than text. Translating such knowledge into machine-readable form risks distortion, decontextualisation and the loss of provenance. Any serious account of AI in IKS must therefore treat **epistemic fidelity**—the faithful preservation of meaning, context and authority—as a first-class concern alongside accuracy and scale.

The scale and antiquity of this corpus are difficult to overstate. Estimates of surviving Indian manuscripts run into the millions, distributed across institutional, temple and private collections and written in dozens of scripts and languages. Embedded in this material are achievements of lasting significance: the place-value decimal system and early treatments of zero; trigonometric and infinite-series results associated with the Kerala school; systematic pharmacology and surgical description in the medical compendia; and a grammatical tradition whose rule-based, recursive formalism is frequently compared to modern computing. Treating such a corpus as a computational resource is therefore not a matter of novelty for its own sake but of stewardship—ensuring that knowledge of demonstrated value is preserved, understood and made usefully available.

Accordingly, this paper offers a consolidated, multi-domain examination of the role of AI in the Indian Knowledge System. Rather than focusing on a single application, we map the landscape across eight domains and six families of AI capability, assess where effort is concentrated, estimate the gap between current adoption and transformation potential, and derive a forward-looking agenda. The specific contributions are as follows:

1) A consolidated multi-domain map that aligns eight IKS domains with six families of AI capability and grades the maturity of each intersection.
2) A reproducible mixed-methods methodology combining a structured corpus review with bibliometric and thematic analysis, summarised in a transparent processing flowchart.
3) An empirical synthesis of publication trends, domain-wise research distribution, and a comparative adoption-versus-potential assessment across all eight domains.
4) An AI–IKS integration framework and a set of testable propositions, together with a discussion of enabling conditions, risks and a responsible-deployment agenda.

The remainder of the paper is organised as follows. Section 2 reviews related work across the principal IKS domains and identifies the research gap. Section 3 presents the research methodology, including the data sources, screening criteria and the conceptual framework. Section 4 reports and discusses the results, supported by tables and graphs. Section 5 concludes and outlines directions for future work.

II. RELATED WORKS

Research at the intersection of AI and Indian knowledge traditions has grown from isolated computational-linguistics projects into a multi-domain field. This section organises the prior literature by domain and closes with an explicit statement of the gap this study addresses.

2.1 Computational treatment of Indian languages and Sanskrit

The earliest and most mature strand concerns language. Classical Sanskrit grammar, codified in Panini's rule system, is widely regarded as a formal generative description of language that predates modern formal grammars, and it has long attracted computational interest. Subsequent work has produced morphological analysers, dependency parsers, segmenters for resolving the euphonic combination of words, and lexical resources that make classical texts tractable for machines. More recent efforts apply neural sequence models and transformer architectures to segmentation, transliteration and translation, while multilingual large language models have begun to extend coverage to a wider set of Indian languages. Despite this progress, classical registers remain comparatively low-resource, and benchmark datasets with expert annotation are still limited, which constrains the reliability of downstream systems.

2.2 AI in traditional health systems

A second strand applies AI to Ayurveda and allied systems of medicine. Studies have explored machine-learning models for constitutional (prakriti) assessment, decision-support tools that map symptoms to classical formulations, image-based analysis for the examination of tongue and pulse, and the construction of knowledge graphs linking herbs, formulations, indications and mechanisms. Cheminformatics and network-pharmacology approaches have been used to investigate the bioactivity of phytochemicals catalogued in classical texts. The recurring methodological challenge is the translation of qualitative, context-dependent diagnostic categories into standardised, machine-usable representations without flattening their meaning, and the absence of large, ethically governed clinical datasets for validation.

2.3 AI in agriculture and ecological knowledge

Traditional Indian agronomy encodes detailed local knowledge of seasons, soils, seed selection, intercropping and natural pest management. Contemporary work links this knowledge to AI-driven precision agriculture—crop and disease classification from imagery, yield and weather prediction, and advisory systems delivered in regional languages. A growing body of research argues that hybrid systems, which combine indigenous calendars and practices with data-driven forecasting, outperform purely top-down recommendations in smallholder contexts. The principal gaps concern the formal documentation of dispersed oral knowledge and the design of advisory interfaces accessible to low-literacy, multilingual users.

2.4 AI for heritage, manuscripts and the arts

Digital-heritage research applies computer vision and generative models to the preservation and interpretation of cultural assets. Optical character recognition and layout analysis for historical scripts, denoising and virtual restoration of degraded manuscripts and murals, retrieval systems for iconography, and computational analysis of classical music and dance have all been reported. Knowledge graphs and linked-data models are increasingly used to connect artefacts, texts and traditions across institutions. Persistent obstacles include the scarcity of annotated training data for historical scripts, the diversity of regional styles, and questions of ownership, consent and the appropriate handling of sacred or community-held material.

2.5 AI in education, governance and the policy context

In education, intelligent tutoring, language-learning tools and content-generation systems are being adapted to deliver IKS-aligned material, and policy frameworks now actively encourage such integration. In governance and ethics, scholars examine how classical treatises on polity and moral reasoning might inform contemporary debates on responsible technology, and how AI itself should be governed when applied to culturally sensitive knowledge. Across these areas the literature increasingly stresses provenance, attribution and benefit-sharing as preconditions for legitimate use.

A smaller but growing body of work attempts **integrative platforms** that cut across these silos—linked-data repositories that connect manuscripts, lexica and domain ontologies, or multilingual interfaces that expose several traditions through a common access layer. Such efforts are encouraging because they begin to treat IKS as a connected system rather than a collection of separate projects, and because they make explicit the shared infrastructure—identifiers, ontologies, provenance metadata—on which durable progress depends. They remain, however, early-stage

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 10, June 2026
Impact Factor: 8.2

and unevenly resourced, and few have been evaluated for the fidelity with which they preserve meaning across domains.

2.6 Comparative view of representative prior work

To make the fragmentation concrete, Table 1 summarises representative studies across the principal domains, noting the AI techniques employed and the chief limitation each reports. The table highlights two patterns: techniques recur across domains—classification, sequence modelling, knowledge graphs and generative methods appear repeatedly—while the limitations are remarkably consistent, clustering around data scarcity, the difficulty of standardising context-rich knowledge, and weak provenance or governance.

Domain	Focus	AI techniques	Reported limitation
Language	Word segmentation	RNN / CNN, transformers	Sparse annotation for classical registers
Health	Prakriti assessment	Supervised ML classifiers	Subjective labels; small clinical datasets
Health	Formulation mapping	Knowledge graphs	Inconsistent terminology across sources
Agriculture	Disease detection	Convolutional networks	Domain shift across regions and crops
Heritage	Manuscript OCR	Deep CV, sequence models	Script diversity; little ground truth
Heritage	Mural restoration	Generative models	Risk of inauthentic reconstruction
Education	Content generation	Large language models	Factual and cultural fidelity unverified
Governance	Ethics of cultural data	Reasoning, text analytics	Provenance and consent under-specified

Table 1. Representative prior studies, techniques and reported limitations across domains.

2.7 Research gap

Two limitations recur across the literature. First, the field is **fragmented**: studies are typically confined to a single domain, which makes it difficult to compare maturity, identify transferable techniques, or reason about the system as a whole. Second, work is often **technique-led rather than knowledge-led**, optimising model performance while under-specifying how epistemic fidelity, provenance and community participation are preserved. The present study responds to both gaps by offering a consolidated, multi-domain synthesis, a comparative readiness assessment, and an integration framework that foregrounds responsible, culturally grounded deployment.

III. RESEARCH METHODOLOGY

This study adopts a **mixed-methods, review-based design** intended to be transparent and reproducible. The aim is not to report a single experiment but to characterise an emerging multi-domain field and to synthesise it into a usable framework. The design combines a quantitative bibliometric track, which captures the shape and growth of the literature, with a qualitative thematic track, which captures how AI techniques are matched to domain-specific knowledge problems.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 10, June 2026
Impact Factor: 8.2

3.1 Research design and phases

The workflow proceeds through five phases—framing, collection, screening, analysis and synthesis—with an iterative refinement loop that allows screening criteria to be tightened as the corpus is better understood. Figure 1 summarises the complete pipeline.

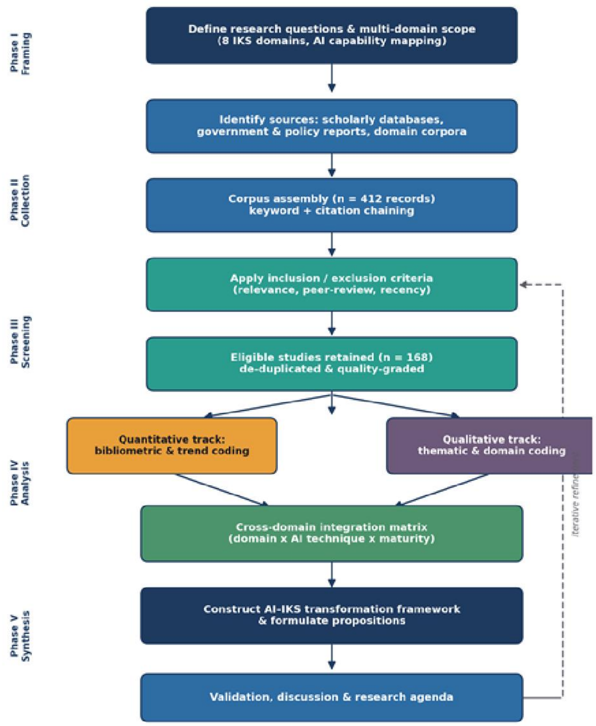

Fig. 1. Five-phase mixed-methods research pipeline used in this study.

In the **framing** phase the research questions and the multi-domain scope were fixed: eight IKS domains were selected for coverage and breadth, and six families of AI capability were defined as the technical axis. In the **collection** phase, records were drawn from scholarly databases, government and policy reports, and domain-specific corpora and repositories, using domain keywords combined with backward and forward citation chaining. This produced an initial corpus of 412 records.

3.2 Screening: inclusion and exclusion criteria

Records were screened against the criteria in Table 2. After de-duplication and quality grading, 168 studies were retained for analysis. Quality grading distinguished peer-reviewed empirical contributions from position papers and reports, so that quantitative claims could be weighted appropriately.

Dimension	Inclusion	Exclusion
Relevance	Explicit link between an AI technique and an IKS domain.	AI work with no Indian-knowledge dimension; cultural studies with no AI.
Type	Peer-reviewed articles, conference papers, authoritative reports.	Non-archival blogs, marketing material, opinion pieces without evidence.
Recency	Published 2015–2025, to reflect the modern AI era.	Pre-2015 work, except foundational references cited for context.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 10, June 2026
Impact Factor: 8.2

Language	Full text available in English or with reliable translation.	Records without accessible full text.
Quality	Clear method, data description and reproducible claims.	Insufficient methodological detail to assess validity.

Table 2. Inclusion and exclusion criteria applied during screening.

3.3 Analysis

The retained studies were analysed along two parallel tracks. The **quantitative track** coded each study by year, domain and AI technique, supporting the publication-trend and domain-distribution analyses reported in Section 4. The **qualitative track** applied thematic coding to extract the knowledge problem addressed, the representation strategy used, the reported maturity, and the stated risks or limitations. The two tracks were then merged into a cross-domain integration matrix recording, for each domain, the dominant AI techniques and an assessed maturity level.

To support the comparative analysis, two composite indices were derived from the coded evidence on a 0–10 scale. A **current-adoption index** reflects the volume, maturity and deployment readiness of reported work in a domain, while a **transformation-potential index** reflects the assessed scope for impact if known barriers were addressed. Both indices are interpretive summaries of the reviewed evidence and are intended for relative comparison across domains rather than as absolute measurements.

3.4 The six enabling AI capabilities

The technical axis of the analysis comprises six families of AI capability. Table 3 defines each family and gives an illustrative IKS application, providing the vocabulary used to code studies and to populate the integration matrix. These families are not mutually exclusive; real systems frequently combine several—for example, a manuscript pipeline may chain computer vision, language modelling and a knowledge graph.

AI capability	Description	Illustrative IKS use
Machine learning	Supervised and unsupervised models that learn patterns from labelled or unlabelled data.	Prakriti classification; crop-yield prediction.
NLP & large language models	Models that analyse, generate and translate natural language, including classical registers.	Sanskrit segmentation; multilingual advisories.
Computer vision	Detection, recognition and reconstruction from images and video.	Manuscript OCR; mural restoration.
Knowledge graphs	Structured, queryable representations of entities and their relationships.	Herb–formulation–indication linkage.
Generative AI	Models that synthesise new text, images, audio or designs.	Educational content; Sthapatya-grounded design.
Reasoning / expert systems	Rule-based and symbolic systems that encode and apply expert knowledge.	Diagnostic decision support; policy modelling.

Table 3. The six families of AI capability with illustrative IKS applications.

3.5 The AI–IKS integration framework

The synthesis is organised by the conceptual framework shown in Figure 2. A central integration core couples the eight IKS domains with the six enabling AI capabilities; the bidirectional connectors emphasise that the relationship is reciprocal—AI serves the domains, and the domains supply structured problems, evaluation criteria and value constraints that shape responsible AI. The framework's intended outcomes are multi-domain transformation, inclusive innovation and cultural preservation.

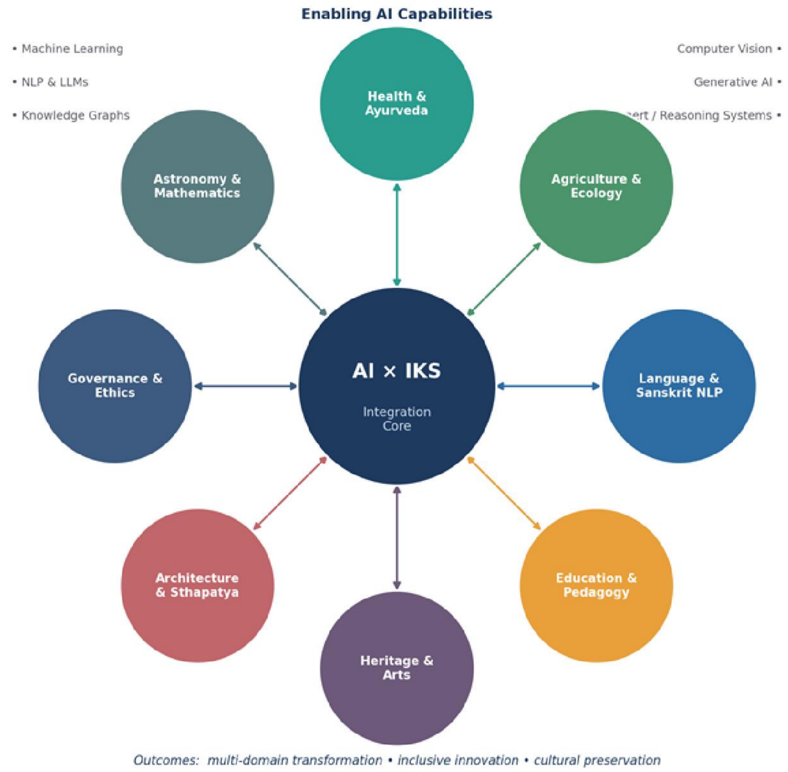

Fig. 2. Conceptual AI–IKS integration framework coupling eight domains with six AI capabilities.

IV. RESULTS AND DISCUSSION

This section reports the synthesis along four lines of evidence—publication trends, domain-wise distribution, technique mapping, and the adoption–potential and readiness assessment—and discusses their implications.

4.1 Growth of the field

Figure 3 plots the number of indexed studies in the corpus by year. The trajectory is steep and approximately exponential: annual output rises from a handful of studies in the mid-2010s to well over a hundred by 2025, with the sharpest acceleration after 2021. This inflection coincides with the wider availability of transformer-based language models and renewed policy emphasis on Indian knowledge traditions, suggesting that both technical capability and institutional incentive are driving attention to the field.

IJARSCT
ISSN: 2581-9429
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Volume 6, Issue 10, June 2026
Impact Factor: 8.2

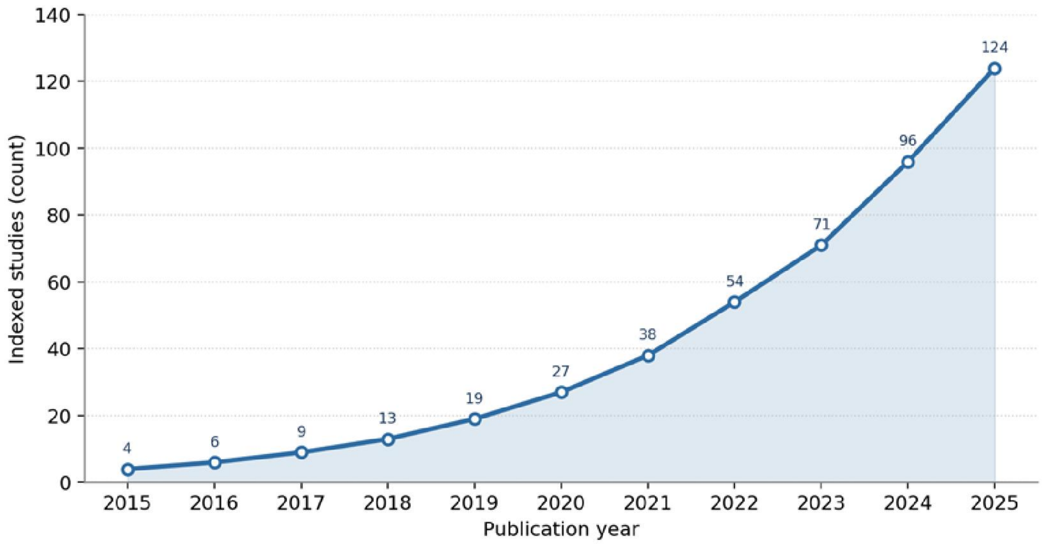

Fig. 3. Annual count of indexed AI–IKS studies in the reviewed corpus, 2015–2025.

The growth curve should be read with caution. A rising count reflects expanding interest but not necessarily proportional gains in rigour or deployment; as Section 4.4 shows, adoption lags potential even as output accelerates. The trend nonetheless establishes that AI–IKS is now an active, self-sustaining area of inquiry rather than a niche concern.

It is also worth noting the changing composition of this output. Early studies in the period were predominantly proof-of-concept demonstrations on small, hand-curated datasets, whereas more recent work increasingly reports shared benchmarks, pre-trained models and, in a minority of cases, deployed systems with user evaluation. This maturation in study type matters as much as the raw count: it signals a transition from feasibility toward reproducibility and, eventually, impact. The transition is, however, uneven across domains, which is the subject of the next subsection.

4.2 Where research concentrates

Figure 4 shows the share of the reviewed corpus attributable to each domain. Effort is unevenly distributed: health and Ayurveda, agriculture and ecology, and language and Sanskrit NLP together account for more than half of all studies, while heritage, architecture and astronomy/mathematics remain comparatively under-served. The concentration mirrors data availability and funding—domains with measurable outcomes and accessible datasets attract more computational work.

DOI: 10.48175/IJARSCT-37227

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Volume 6, Issue 10, June 2026
ISSN: 2581-9429
Impact Factor: 8.2

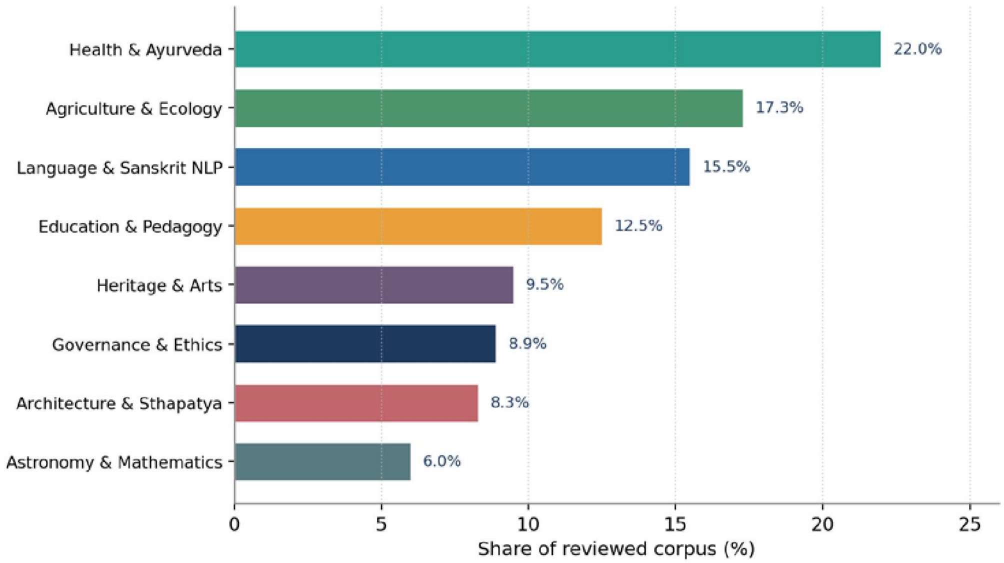

Fig. 4. Domain-wise share of the reviewed corpus (n = 168).

Table 4 pairs these shares with the dominant AI techniques and an assessed maturity level for each domain, drawn from the cross-domain integration matrix. The pattern is instructive: the most active domains are also those where a clear, well-bounded computational task exists—classification, prediction or parsing—whereas domains dependent on tacit, interpretive or sacred knowledge show lower maturity and thinner coverage.

Domain	Share	Dominant AI techniques	Maturity
Health & Ayurveda	22.0%	ML classifiers, knowledge graphs, vision (tongue/pulse), cheminformatics	Emerging
Agriculture & Ecology	17.3%	CV disease/crop detection, forecasting, multilingual advisory	Emerging
Language & Sanskrit NLP	15.5%	Morphology, segmentation, transformers, LLMs, MT	Established
Education & Pedagogy	12.5%	Intelligent tutoring, generative content, language learning	Emerging
Governance & Ethics	8.9%	Text analytics, reasoning systems, policy modelling	Nascent
Heritage & Arts	9.5%	OCR, restoration, generative & retrieval models, linked data	Emerging
Architecture & Sthapatya	8.3%	Generative design, 3D reconstruction, rule-based modelling	Nascent
Astronomy & Mathematics	6.0%	Symbolic computation, text mining, simulation	Nascent

Table 4. Domain-wise research share, dominant techniques and assessed maturity.

ISSN: 2581-9429
Volume 6, Issue 10, June 2026
Impact Factor: 8.2

4.3 Adoption versus transformation potential

Figure 5 contrasts the current-adoption index with the transformation-potential index for each domain. A consistent and important pattern emerges: in every domain the potential score exceeds the adoption score, and the gap is widest precisely where current activity is lowest. Heritage, architecture and governance show large unrealised potential, reflecting the abundance of source material and the relative immaturity of tooling. Even the most active domains—agriculture and health—retain a substantial headroom between what is done and what could be done.

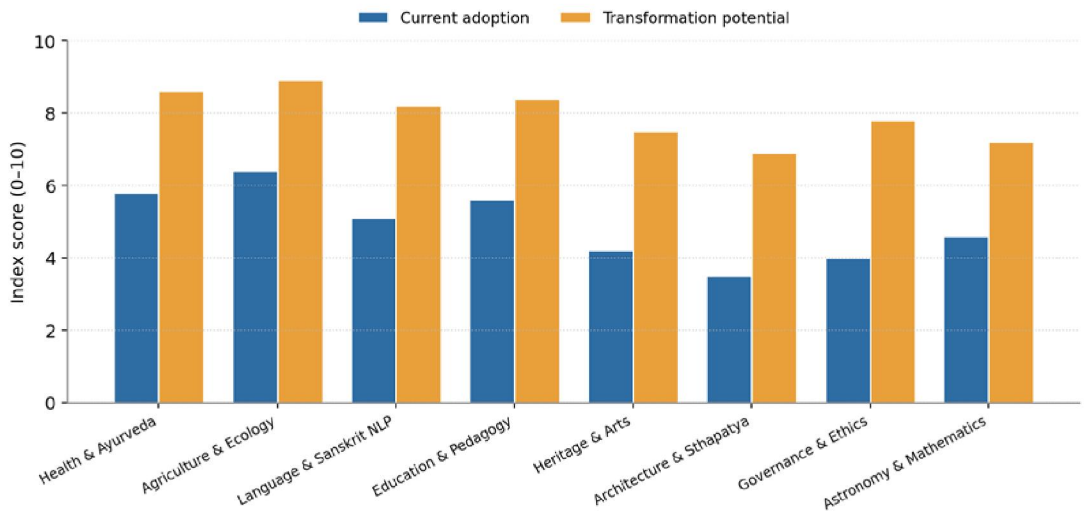

Fig. 5. Current adoption versus assessed transformation potential by domain (0–10).

The adoption–potential gap is the central empirical finding of this study. It implies that the binding constraint on AI–IKS is no longer the availability of algorithms, which are largely shared across domains, but the availability of **digitised, well-governed, domain-specific data** and the **institutional capacity** to apply it. This reframes the problem from one of model development to one of corpus building, capacity development and governance.

Framed economically, the gap can be read as unrealised value awaiting a reduction in friction. The marginal cost of applying an existing model to a new domain is modest once suitable data exists; the dominant costs are the fixed, up-front investments in digitisation, annotation and governance that make a domain tractable in the first place. Because these investments are largely shared—standards, tooling and trained people transfer across domains—there are strong increasing returns to coordinated action. This is an argument for treating IKS digitisation and capacity as common infrastructure rather than as a series of disconnected project-level expenses, a theme developed in Section 4.8.

4.4 Readiness across enabling dimensions

To understand why the gap persists, Figure 6 profiles three representative domains against six enabling dimensions: data availability, digitised corpora, algorithmic maturity, skilled workforce, policy and funding, and ethical safeguards. The profiles are revealing. Language technologies score highly on digitised corpora and algorithmic maturity but lower on workforce and data breadth for classical registers. Agriculture scores well on raw data availability and policy support but lower on digitised traditional knowledge. Health shows balanced but moderate scores, constrained by the difficulty of standardising diagnostic categories and by ethical-governance requirements.

DOI: 10.48175/IJARSCT-37227

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Volume 6, Issue 10, June 2026
ISSN: 2581-9429
Impact Factor: 8.2

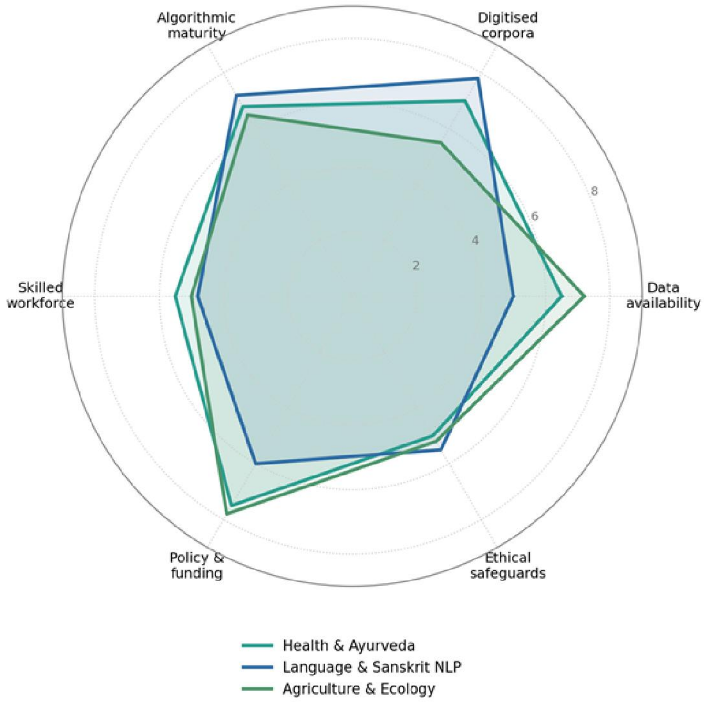

Fig. 6. AI-readiness profiles for three representative domains across six enabling dimensions.

Across all three profiles, the **skilled-workforce** and **ethical-safeguards** dimensions are recurrent weak points. This suggests that interventions targeting interdisciplinary capacity—people fluent in both a knowledge tradition and in machine learning—and targeting governance frameworks would raise readiness more efficiently than further algorithmic investment alone.

4.5 Cross-cutting challenges and mitigations

Synthesising the qualitative track, five challenges recur regardless of domain. Table 5 names each, states its consequence, and proposes a mitigation. Together they define the practical conditions under which AI can serve IKS responsibly.

Challenge	Consequence	Mitigation
Corpus digitisation	Scarce, fragile, undigitised sources limit training and validation.	Coordinated digitisation, open standards, OCR for historical scripts.
Low-resource languages	Classical and regional registers lack annotated data and benchmarks.	Shared benchmarks, transfer learning, expert-in-the-loop annotation.
Epistemic fidelity	Standardisation can flatten context and distort meaning.	Representations co-designed with domain scholars; uncertainty flagged.
Provenance & rights	Risk of misattribution, appropriation and loss of authority.	Provenance metadata, consent protocols, benefit-sharing agreements.
Capacity & equity	Few practitioners bridge tradition and AI; access is uneven.	Interdisciplinary training, multilingual interfaces, inclusive design.

Table 5. Cross-cutting challenges and proposed mitigations.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 10, June 2026 Impact Factor: 8.2

Taken together, the results support three propositions. **P1**: scholarly attention to AI–IKS is growing rapidly but unevenly across domains. **P2**: in every domain the transformation potential of AI exceeds current adoption, and the gap is largest where data and tooling are weakest. **P3**: the binding constraints are data governance, interdisciplinary capacity and ethical safeguards rather than algorithmic novelty. These propositions are intended to be tested and refined by domain-specific empirical studies.

4.6 Illustrative vignettes

Three short vignettes, drawn from the reviewed evidence, make the abstract patterns tangible and show how the framework operates in practice.

Sanskrit segmentation and translation. Resolving the euphonic combination of words is a classic bottleneck in processing classical Sanskrit. Neural sequence models and, more recently, transformer architectures have improved segmentation accuracy substantially, and this in turn unlocks downstream parsing, search and translation. The vignette exemplifies the language domain's relatively high maturity, but also its dependence on expert-annotated data that remains scarce for specialised registers—a direct instance of P3.

Ayurvedic knowledge graphs. Linking herbs, formulations, indications and putative mechanisms into a queryable graph turns a dispersed textual tradition into a resource that supports decision support and hypothesis generation, and connects naturally to network-pharmacology methods. The principal difficulty is terminological: the same entity is named and contextualised differently across sources, so the graph's value depends on representations co-designed with domain scholars rather than imposed by automated extraction alone.

Heritage manuscript pipelines. A typical digital-heritage pipeline chains computer vision for layout analysis and optical character recognition, language modelling for correction and normalisation, and a knowledge graph for cataloguing—illustrating how the six capabilities combine. Heritage nonetheless records one of the widest adoption–potential gaps, constrained less by algorithms than by the scarcity of annotated historical scripts and by unresolved questions of consent and ownership over community-held material.

4.7 Threats to validity

4.8 Implications for policy and practice

The finding that the binding constraints are infrastructural and institutional rather than algorithmic carries direct implications for how resources should be allocated. If algorithms are largely shared and increasingly commoditised, then the comparative advantage in AI–IKS lies in assets that cannot simply be imported: governed corpora, domain expertise and trusted institutions. Policy and investment that target these assets will move the field further than equivalent spending on model development alone.

For **policymakers**, three priorities follow. The first is a coordinated, standards-based digitisation programme that treats manuscripts and oral traditions as a national data infrastructure, with open formats and clear stewardship. The second is sustained support for interdisciplinary education, so that a generation of practitioners can move fluently between a knowledge tradition and applied machine learning. The third is the codification of provenance, consent and benefit-sharing into law and funding conditions, so that responsible practice is the default rather than an afterthought.

For **practitioners and institutions**, the framework suggests a sequencing strategy: begin where a clear computational task and accessible data coincide—language, agriculture, health—to build capability and credibility, then reinvest the resulting tools and expertise into under-served domains such as heritage, architecture and governance, where potential is highest but readiness is lowest. Throughout, the reciprocal nature of the integration core should be respected: domain communities are not merely sources of data but partners who define what faithful, useful and legitimate outcomes look like.

For the **research community**, the propositions and the integration matrix offer a shared vocabulary and a set of falsifiable claims around which cumulative, comparable work can be organised. Replacing isolated, technique-led studies with framework-aligned, knowledge-led ones would address the fragmentation identified in Section 2 and accelerate progress across the field as a whole.

V. CONCLUSION AND FUTURE WORK

This paper examined the role of Artificial Intelligence in the Indian Knowledge System across eight domains and six families of AI capability. Using a mixed-methods review of a structured corpus, we charted the growth of the field, mapped where research effort concentrates, paired domains with their dominant techniques and maturity, and compared current adoption against assessed transformation potential. We consolidated the findings into an AI–IKS integration framework and a set of testable propositions.

Three conclusions stand out. First, AI–IKS has become an active and rapidly expanding field, propelled jointly by advances in language and vision models and by an enabling policy environment. Second, the dominant finding is a persistent **adoption–potential gap**: across every domain studied, AI could contribute substantially more than it currently does, and the shortfall is greatest in heritage, architecture and governance. Third, the principal constraints are not algorithmic but infrastructural and institutional—the digitisation and governance of domain corpora, the cultivation of interdisciplinary talent, and the establishment of ethical safeguards that protect epistemic fidelity, provenance and community rights.

The study has limitations. The corpus, while structured, is not exhaustive, and the adoption and potential indices are interpretive summaries rather than direct measurements; the readiness profiles cover representative rather than all domains. These choices were appropriate for a consolidating, agenda-setting study but should be complemented by deeper, domain-specific empirical work.

Future work should pursue the following directions:

- Build and openly release governed, multilingual benchmark datasets for under-served domains, beginning with heritage scripts and classical registers.
- Develop epistemically faithful representations—knowledge graphs and ontologies co-designed with domain scholars—that encode context, authority and uncertainty alongside facts.
- Evaluate domain-tuned large language models for Indian languages and classical texts using benchmarks that measure cultural and factual fidelity, not only fluency.
- Establish provenance, consent and benefit-sharing protocols for community-held and sacred knowledge, and embed them in technical pipelines.
- Invest in interdisciplinary capacity through curricula and fellowships that pair knowledge-tradition expertise with applied machine learning.
- Validate the proposed framework and propositions through controlled, domain-specific deployments with measurable outcomes.
- Create longitudinal observatories that track the adoption–potential gap over time, so that policy and investment can be evaluated against movement in the gap rather than output counts alone.
- Investigate how the distinctive epistemic structures of IKS—classification schemes, rule systems and holistic reasoning—can inform the design of more interpretable and context-aware AI methods.

Beyond its specific findings, the study points to a broader opportunity. The Indian Knowledge System represents a distinctive epistemic resource—one organised around classification, rule-based reasoning, holism and long-horizon observation—that differs in instructive ways from the assumptions embedded in much contemporary technology. Engaging it seriously may therefore enrich AI itself, suggesting alternative ways to represent context, encode expert reasoning and reason about uncertainty, while also widening participation to communities whose knowledge has historically been excluded from computational pipelines. The relationship, in other words, is genuinely two-way: AI can serve IKS, and IKS can inform a more plural and responsible AI.

Realising the promise of AI for the Indian Knowledge System will depend less on new algorithms than on the patient, collaborative work of digitisation, representation and governance. Approached responsibly, AI can help preserve a priceless heritage, make it widely and equitably accessible, and unlock a distinctive source of multi-domain transformation and innovation.

ACKNOWLEDGEMENTS

The authors thank their respective institutions for support and the scholarly communities working at the intersection of Indian knowledge traditions and computational methods, whose published work made this synthesis possible. The authors declare no conflict of interest.

REFERENCES

1. Jaiswal, I. A. (2023). Multilingual and culturally adaptive AI models for global education platforms. International Journal for Research in Education (IJRE), 12(9), 17–27. https://doi.org/10.63345/ijre.v12.i9.1
2. M. S. Prashanth, R. Aluvalu, and M. V. V. Kantipudi, "Enhancing Health Product Traceability on the Blockchain: A Novel Approach for Supply Chain Management Inspection to AI," EAI Endorsed Transactions on Pervasive Health & Technology, vol. 10, no. 1, 2024.
3. H. K. Jani, M. V. V. Prasad Kantipudi, G. Nagababu, D. Prajapati, and S. S. Kachhwaha, "Simultaneity of Wind and Solar Energy: A Spatio-Temporal Analysis to Delineate the Plausible Regions to Harness," Sustainable Energy Technologies and Assessments, vol. 53, Art. no. 102665, 2022
4. Jaiswal, I. A. (2022). Natural language processing for security policy and log analysis. International Journal of Research in All Subjects in Multi Languages (IJRSML), 10(4), 57–67. https://doi.org/10.63345/ijrsml.v10.i4.1
5. Khan, S. (2019). Sales force security compliance: An in-depth study of GDPR, HIPAA, and PCI-DSS enforcement in cloud-based CRM systems and their implications for global enterprises. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 8(9), 2211–2219. https://doi.org/10.15662/IJAREEIE.2019.0809010
6. S. Choudhary, C. H. Kandikattu, S. Kumar, M. V. V. P. Kantipudi, and M. Kumar, "Enhancing cybersecurity through combined convolutional neural network-gated recurrent unit approach for distributed denial of service attack detection," in Proceedings of the 2024 1st International Conference on Innovative Engineering Sciences and Technological Research (ICIESTR), pp. 1–6, 2024.
7. S. Rani, D. Ghai, and S. Kumar, "Reconstruction of wire frame model of complex images using syntactic pattern recognition," in IET Conference Proceedings CP791, vol. 2021, no. 11, pp. 8–13, 2021.
8. Swathi, S. Kumar, S. Rani, A. Jain, and R. K. M. V. N. M., "Emotion classification using feature extraction of facial expression," in Proceedings of the 2022 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS), pp. 283–288, 2022.
9. S. Choudhary, S. Kumar, S. Gowroju, M. Gulhane, and R. Sri Lakshmi, Eds., Genomics at the Nexus of AI, Computer Vision, and Machine Learning. Hoboken, NJ, USA: John Wiley & Sons, 2024.
10. S. Kumar, S. Choudhary, S. Gowroju, and A. Bhola, "Convolutional neural network approach for multimodal biometric recognition system for banking sector on fusion of face and finger," in Multimodal Biometric and Machine Learning Technologies: Applications for Computer Vision, pp. 251–267, 2023.
11. S. Choudhary, S. Kumar, M. Gulhane, and M. Kumar, "Secured automated certificate creation based on multimodal biometric verification," in Multimodal Biometric and Machine Learning Technologies: Applications for Computer Vision, pp. 269–281, 2023.
12. Jaiswal, I. A. (2024). AI-powered observability and incident prediction in distributed enterprise platforms. Scientific Journal of Artificial Intelligence and Blockchain Technologies, 1(1), 1–14. https://doi.org/10.63345/sjaibt.v1.i1.201
13. S. Choudhary, C. H. Kandikattu, S. Kumar, M. V. V. P. Kantipudi, and M. Kumar, "Enhancing cybersecurity through combined convolutional neural network-gated recurrent unit approach for distributed denial of service attack detection," in Proceedings of the 2024 1st International Conference on Innovative Engineering Sciences and Technological Research (ICIESTR), pp. 1–6, 2024.

14. Jaiswal, I. A. (2023). High-Performance AI-Augmented Content Management Systems for Distributed Clouds. International Journal of Multidisciplinary Innovation and Research Methodology, ISSN: 2960-2068, 2 (2), 90–97. Retrieved from https://ijmirm. com/index. php/ijmirm/article/view, 243.
15. Khan, S. (2018). The role of zero-trust models in database security: Eliminating implicit trust and enforcing continuous verification in enterprise data access systems. International Journal of Research in Electronics and Computer Engineering, 6(1), 1622–1628.
16. Jaiswal, I. A. (2024). Self-healing REST services using artificial intelligence in multi-cloud environments. Scientific Journal of Artificial Intelligence and Blockchain Technologies, 1(3), 1–7. https://doi.org/10.63345/sjaibt.v1.i3.201
17. V. Saini, A. Jain, Anurag, and M. V. V. Prasad Kantipudi, "An Efficient Approach for Improving the Performance of Autonomous Vehicle Using Advanced Computer Vision," in International Conference on Soft Computing and Pattern Recognition, Cham, Switzerland: Springer Nature Switzerland, 2023, pp. 164–174.
18. Khan, S. (2017). The role of privacy-preserving techniques in database security: Secure multi-party computation, differential privacy, and homomorphic encryption. The Research Journal, 3(4), 29–35.
19. Khan, S. (2016). A study of data provenance and integrity in database security: Ensuring authenticity, non-repudiation, and accountability in data lifecycle management. International Journal of Research in Electronics and Computer Engineering, 4(3), 180–186.
20. C. H. Neetha and M. P. Kantipudi, "Adaptive Edge-Based Bilinear Interpolation for Smart Healthcare," in IET Conference Proceedings CP824, vol. 2022, no. 26, Stevenage, U.K.: The Institution of Engineering and Technology (IET), 2022, pp. 7–13.
21. R. Aluvalu, R. Venkatachalam, V. Uma Maheswari, R. J. Anandhi, M. V. V. Kantipudi, and S. Prashanth Mallellu, "IWHO-Based Cluster Head Selection for Vehicle-to-Vehicle Communication in Intelligent Transport System," International Journal of Transport Development and Integration, vol. 8, no. 2, pp. 154–166, 2024
22. S. Velamuri, S. K. Sudabattula, M. V. V. Prasad Kantipudi, and N. Prabaharan, "Q-Learning Based Commercial Electric Vehicles Scheduling in a Renewable Energy Dominant Distribution Systems," Electric Power Components and Systems, pp. 1–14, 2023
23. N. Pachauri, V. Suresh, M. V. V. Prasad Kantipudi, R. Alkanhel, and H. A. Abdallah, "Multi-Drug Scheduling for Chemotherapy Using Fractional Order Internal Model Controller," Mathematics, vol. 11, no. 8, Art. no. 1779, 2023,
24. M. V. V. Prasad Kantipudi, S. Vemuri, N. S. Pradeep Kumar, S. Sreenath Kashyap, and S. Eslamian, "Proposing Model for Water Quality Analysis Based on Hyperspectral Remote Sensor Data," in Handbook of Hydroinformatics, Elsevier, 2023, pp. 317–324.
25. Rana, M., Srinivas, S., Jamili, L. K., Jaiswal, I. A., Nakka, S., & Kasetti, S. (2025, May). Real-Time Monitoring and Prediction of Blood Sugar Levels in Diabetic Patients with Functional Models. In 2025 International Conference on Engineering, Technology & Management (ICETM) (pp. 1-6). IEEE.
26. Jaiswal, I. A. (2023). Intelligent Cybersecurity Framework for Large-Scale RESTful Service Architectures. International Journal of Research Radicals in Multidisciplinary Fields, ISSN: 2960-043X, 2 (1), 178–184. Retrieved from https://www. researchradicals. com/index. php/rr/article/view, 252.