



Non-Western Logic Traditions for Explainable AI: A Survey and Research Agenda

Abhinaba Basu^{1,2} and Vikas Mittal¹

¹ National Institute of Electronics and Information Technology, India

² Indian Institute of Information Technology, Allahabad, India
mail@abhinaba.com, {abhinaba.basu,vikas}@nielit.gov.in

Abstract. Explainable AI (XAI) research predominantly relies on Western logic traditions-Aristotelian syllogisms, Toulmin's argumentation schemes, and formal symbolic logic. However, non-Western cultures have developed rich epistemological frameworks over millennia that offer complementary perspectives for interpretable AI systems, particularly for modern machine learning models and large language models.

We present, to our knowledge, the first dedicated cross-tradition survey of non-Western logic traditions explicitly oriented toward XAI applications, covering traditions from India (Nyaya, Buddhist pramana, Jain logic, Yoga epistemology), East Asia (Neo-Confucian li-qi, Japanese Bud-dhist innen, Chinese Mohism), Africa (Ubuntu philosophy, Ethiopian rationalism), and Islamic scholarship (Kalam, Falsafa, Sufi epistemology). For each tradition, we identify (1) core epistemological principles, (2) structural reasoning patterns, (3) potential XAI applications, and (4) any existing computational implementations.

Key contributions: (i) Concrete illustrative example contrasting West-ern SHAP explanations with Nyaya-structured, Jain syadvada, and Ubuntu-relational framings for medical diagnosis; (ii) Comprehensive summary table mapping current XAI methods to proposed innovations from twelve traditions, serving as an actionable practitioner reference; (iii) Detailed research agenda with specific tasks, datasets, and success criteria span-ning 1-10 year horizons; and (iv) Critical gap analysis showing that most traditions currently lack computational implementations and empirical validation. We review earlier work on Nyaya-inspired systems that reach about 0.91 ERASER sufficiency while still keeping model accuracy high. We also show how ideas from Buddhist apoha, Jain seven-fold predication and Ubuntu-style relational ethics could guide new methods for counterfactual examples, handling uncertainty from several viewpoints, and measuring relational fairness. Overall, this survey aims to move XAI beyond a purely Western way of thinking and to support cross-cultural discussion on how we design and evaluate AI explanations.

Keywords: Explainable AI · Non-Western Logic · Nyaya · Buddhist Logic · Ubuntu · Decolonial AI · Cultural AI · Interpretability.

(Article history: Selected papers from NIELIT's International Conference on Communication, Electronics and Digital Technologies (NICE-DT'26), 29-30Jan, 2026, Gorakhpur, India)

1 Introduction

1.1 Motivation: The Western-Centric Nature of XAI

Modern XAI has made strong progress with methods such as LIME [1], SHAP [2], attention mechanisms [3], and chain-of-thought prompting [4]. Most of this work is based on Western ideas. LIME and SHAP build on game theory and 20th-century statistics. Attention mechanisms are often linked to Western cognitive psychology, and many explanation styles follow Toulmin's model of argumenta-tion [5]. These methods are useful, but this narrow focus leaves out rich ways of thinking about knowledge that have been developed in other cultures over thousands of years.

1.2 The Case for Epistemological Diversity

Non-Western logic traditions offer complementary advantages for XAI. **Structural diversity** provides alternative reasoning patterns: Nyaya's five-part arguments explicitly chain claim, reason, universal principle, application, and conclusion [6]; Buddhist apoha (definition by negation) naturally supports counterfactual explanations [7]; Ubuntu's relational ontology emphasises communal context over isolated features [8]. **Multiple evidence types** are recognized beyond empirical observation: Nyaya distinguishes perception, inference, comparison, and testimony [6]; Islamic Kalam separates rational, textual, and con-sensual evidence [9]; Confucian li-qi dualism distinguishes principle from material manifestation [10]. **Explicit uncertainty handling** appears in Buddhist catuṣkoṭi (four-valued logic beyond binary true/false) [11], Jain syadvada (context-dependent truth) [12], and Mohist reproducibility standards [13].

Our contributions: (i) First systematic survey of non-Western logic traditions for XAI, covering Indian (Nyaya, Buddhist, Jain, Yoga), East Asian (Neo-Confucian, Japanese Buddhist, Mohist), African (Ubuntu, Ethiopian), and Islamic (Kalam, Falsafa, Sufi) traditions; (ii) Comparative analysis with comprehensive summary tables mapping current XAI methods to proposed innovations;

(iii) Critical assessment of existing computational implementations and research gaps; (iv) Concrete research agenda with specific tasks, methodologies, and success criteria spanning 1-10 year horizons.

1.3 Review Methodology

This narrative scoping review searched major AI/ML venues, philosophical databases, and Google Scholar using tradition-specific keyword combinations (e.g., "Nyaya AND explainable AI"), including both primary philosophical texts and secondary AI/XAI literature.

2 Indian Logic Traditions

India developed sophisticated logical and epistemological systems over more than two millennia. We examine four major traditions with potential XAI applications.

2.1 Nyaya: Classical Indian Logic

Historical Context and Core Principles Nyaya (literally "method" or "rule") is one of six orthodox schools of Indian philosophy, with its foundational text, the Nyaya-sutra, traditionally attributed to Akṣhapada Gautama (c. 2nd century CE) [6]. The tradition reached its classical formulation in the work of Udyotakara (6th century CE) and Vacaspati Miśra (9th century CE).

Core epistemological framework: Nyaya identifies **four pramanas** (means of valid knowledge):

1. **Pratyakṣa (perception):** Direct sensory knowledge, with extensive analysis of perception conditions and error sources.
2. **Anumāna (inference):** Logical deduction based on invariable concomitance (vyāpti).
3. **Upamāna (comparison/analogy):** Knowledge through similarity to known objects.
4. **Śabda (testimony):** Reliable verbal testimony from trustworthy sources.

The Five-Part Argument (Pancavayava): Nyaya formalizes inference through five components [6]:

1. **Pratijñā (proposition):** The claim to be established (e.g., "There is fire on the mountain").
2. **Hetu (reason):** The evidence supporting the claim (e.g., "Because there is smoke").
3. **Udaharana (universal example):** The general principle with supporting case (e.g., "Wherever there is smoke, there is fire-like in a kitchen").
4. **Upanaya (application):** Applying the principle to the current case (e.g., "This mountain has smoke").
5. **Nigamaṇa (conclusion):** Restating the proven claim (e.g., "Therefore, there is fire on the mountain").

Vyāpti (Invariable Concomitance): The concept of vyāpti is central to Nyaya inference—it represents the invariable relation between hetu (reason) and sādhyā (what is to be proved). This is analogous to but distinct from material implication in Western logic [16]. **Hetvabhāsa (Fallacies):** Nyaya developed sophisticated fallacy theory, identifying five types of defective reasoning: Asiddha (unestablished reason), Savyabhicara (irregular reason), Satpratipakṣa (contradicted reason), Badhita (sublated reason), and Viruddha (contrary reason). **XAI Applications: Structured Explanations** Nyaya's structured argument form maps naturally onto explanation generation:

Table 1. Mapping Nyaya Components to XAI

Nyaya Component	XAI Equivalent	Implementation
Pratijna	Predicted class/answer	Model output
Hetu	Supporting evidence	Attribution methods (IG, SHAP)
Udaharana	General principle	Knowledge graph relations, rules
Upanaya	Application	Linking evidence to claim
Nigamana	Conclusion	Explanation summary

Existing computational implementation: The most developed application is **Sutra-PCM** [15], which architecturally integrates Nyaya structure into a commonsense reasoning system. It uses a frozen sentence encoder (MPNet) and a trainable reasoning module. Explanation generation involves multi-method attribution to identify hetu sentences, with ConceptNet providing udaharana hints. Validation using ERASER metrics shows sufficiency ≈ 0.91 , indicating explanations retain 91% of prediction probability.

Research Gaps and Opportunities Despite Sutra-PCM's success, significant gaps remain:

- **Gap 1: Lexical vs. Structural.** Current implementations use lexical markers as proxies. Needed: Structural probes detecting inference patterns.
- **Gap 2: Monolingual Limitation.** Existing work focuses on English. Needed: Multilingual validation across Sanskrit, Hindi, English.
- **Gap 3: Prescriptive vs. Descriptive.** Unknown whether standard transformers naturally encode Nyaya-compatible patterns. Needed: Probe-based analysis.
- **Gap 4: Hetvabhasa Detection.** No work on automatically identifying fallacies. Needed: Fallacy-aware explanation systems.
- **Gap 5: Human Validation.** No studies with Nyaya scholars. Needed: Interdisciplinary collaboration.

2.2 Illustrative Example: Medical Diagnosis Across Traditions

To demonstrate differences in approach, consider a medical AI diagnosing Type 2 Diabetes for a 52-year-old patient with: fasting glucose = 142 mg/dL, BMI = 31.2, HbA1c = 7.1%, family history = positive. Model prediction: **Type 2 Diabetes (0.87 confidence)**.

Key insight: Each serves different communicative goals. Western SHAP is algorithmically faithful; Nyaya shows reasoning chain; Jain acknowledges multiple perspectives; Ubuntu contextualizes socially. A mature XAI system might offer multiple modes depending on audience.

2.3 Buddhist Pramana Theory

Historical Context and Core Principles Buddhist epistemology, developed by Dignaga (c. 480-540 CE) and Dharmakirti (c. 600-660 CE), recognizes two valid means of knowledge: **Pratyaksha** (perception) and **Anumana** (inference) [17]. Key distinction: *svarthanumana* (inference for oneself) vs. *pararthanu-mana* (inference for others)—relevant to XAI where explanations target external audiences.

Apoha (Exclusion): Dharmakirti's theory defines concepts through exclusion rather than positive essence [7]. For example, "cow" means "not non-cow." This maps naturally onto counterfactual explanations [18]: explaining a prediction by showing minimal changes that would alter it.

Catuskoti (Four-Valued Logic): Buddhist logic uses four values [11]:

(1) Is, (2) Is not, (3) Both, (4) Neither. This framework handles paradoxical or context-dependent statements common in AI applications, mapping to: high-confidence positive, high-confidence negative, ambiguous cases, and out-of-distribution samples.

Table 2. Comparison of Explanation Approaches for Diabetes Diagnosis

Approach	Explanation Style
Western SHAP	Feature contributions: HbA1c +0.31 (strongest), Glucose +0.22, BMI +0.18, Family history +0.12. Mathematically rigorous but lacks narrative.
Nyaya 5-Part	Pratijna: Patient has Type 2 Diabetes. Hetu: Elevated HbA1c (7.1%), glucose (142), BMI (31.2). Udaharana: Medical principle: HbA1c $\geq 6.5\%$, glucose ≥ 126 , BMI > 30 indicate diabetes (94% in training cohort). Upanaya: This patient exceeds all thresholds. Nigamana: Therefore, diabetes with 87% confidence. Explicit reasoning chain, pedagogically valuable.
Jain Syadvada	Syad asti: From glycemic perspective, patient <i>has</i> diabetes. Syad asti ca nasti ca: Shows clear dysfunction (is diabetic) but borderline symptoms suggest early-stage (not fully progressed)—recommend confirmatory test. Syad avaktavya: Genetic/long-term risk incompletely captured—additional profiling needed. Acknowledges uncertainty, encourages validation.
Ubuntu Relational	Diagnosis affects family (children have 40% higher risk—recommend screening) and community (third case this month in neighborhood suggests environmental factors). Context: community-level intervention needed, not just individual treatment. Reframes AI as socially embedded.

XAI Applications and Research Gaps Apoha for Counterfactuals: Definition-by-exclusion aligns with counterfactual methods asking “What minimal change would flip the prediction?” **Catuskoti for Uncertainty:** Four-valued logic provides richer uncertainty than binary classification, mapping to ensemble disagreement. **Key gaps:** No computational implementations of apoha-based counterfactual generation exist; catuskoti lacks ML formalization; architectural implications of pratyaksha/anumana distinction unexplored.

2.4 Jain Logic and Syadvada

Core Principles Jainism developed **anekantavada** (many-sidedness) and **syad-vada** (conditional predication), recognizing truth as multifaceted and context-dependent [12]. The **saptabhengi** (seven-fold predication) offers seven qualified assertions: (1) *Syad asti* (maybe, it is), (2) *Syan nasti* (maybe, it is not), (3) *Syad asti nasti* (both), (4) *Syad avaktavya* (indescribable), plus combinations. The term *syad* emphasizes conditionality and perspective-dependence.

XAI Applications and Research Gaps Multi-view learning: Saptabhengi fits well with ensemble methods, where each model captures only part of the truth. Together, they give explanations that admit several valid ways to see the same case.

Qualified statements: Instead of giving a single probability, we can qualify predictions by perspective. For example: from feature set A the case looks positive, from feature set B it looks negative, and if we consider both together the result is ambiguous.

Indescribability for OOD: The *avaktavya* category handles out-of-distribution samples where models should abstain. **Key gaps:** No computational implementations exist; syadvada’s multi-perspective approach could inform multi-stakeholder fairness frameworks.

2.5 Yoga Epistemology and Atmapariksha

Yoga philosophy emphasizes **atmapariksha** (self-examination) and **sakshin** (witness consciousness) [19]. This inspires “Self-Examination” frameworks for AI, where models examine their own reasoning for biases before answering, similar to Constitutional AI [20] and self-refinement [21].



3 East Asian Logic Traditions

3.1 Neo-Confucian Logic

Core Principles Neo-Confucianism, particularly Zhu Xi (1130-1200 CE), developed **li-qi** theory [10]: *Li* (principle, rational structure) vs. *Qi* (material force, substance). Another key concept is **gewuzhizhi** (investigating things to extend knowledge): empirical investigation combined with pattern recognition.

XAI Applications Li-Qi for Neural Networks: *Li* maps to learned weights/patterns/abstract representations; *Qi* maps to specific activations/concrete inputs. Explanations distinguish “principle-level” (how model generally works) vs. “instance-level” (why this specific prediction). **Ethical Dimensions:** Confucian *ren* (benevolence), *li* (propriety), and *yi* (righteousness) inform explanation design for human welfare and contextual appropriateness.

3.2 Japanese Buddhist Logic (Innen)

Core Principles and XAI Applications Japanese Buddhism emphasizes **in-nen** (dependent origination): every phenomenon arises through causes and conditions [22]. Nothing exists independently; all is interconnected (*pratityasamutpada*, *shunyata*/emptiness).

XAI Applications: Deep networks exemplify dependent origination—each neuron depends on previous layers. Explanations should trace causal chains rather than isolated features (holistic vs. atomistic). Emptiness implies models lack “independent essence” (defined by training data/architecture), motivating robustness testing.

3.3 Chinese Mohist Logic

Core Principles and XAI Applications Mohism, founded by Mozi (c. 470-391 BCE), developed rigorous argumentation and empirical verification [13]. **Three Standards (San Biao)** for claim validity: (1) **Ben** (foundation): historical precedent, authority; (2) **Yuan** (source): empirical verification, observation; (3) **Yong** (application): practical utility, societal benefit. Valid claims satisfy all three.

XAI Meta-Evaluation: The three standards assess explanation techniques: Does it have theoretical grounding (ben)? Empirically valid, reproducible results (yuan)? Practical value to users (yong)? This encourages holistic evaluation beyond accuracy metrics, combining theoretical soundness, empirical validation, and real-world usefulness.

4 African Logic Traditions

4.1 Ubuntu Philosophy

Core Principles and XAI Applications Ubuntu (“I am because we are”) emphasizes relationality, communalism, and interdependence [8]. Identity emerges from community connections; actions affect the whole community; knowledge emerges through dialogue.

XAI Reframing: Traditional XAI asks “Why this prediction for this individual?” Ubuntu reframes: “How does this affect the community? What are relational consequences?” **Relational Fairness:** Move beyond individual-level metrics to analyze family/network impacts, community cohesion disruption, and clustered denials (e.g., credit scoring affecting intergenerational wealth transfer). **Participatory XAI:** Affected communities help design satisfactory explanations.

4.2 Ethiopian Philosophy

Core Principles and XAI Applications Zera Yacob (1599-1692), Ethiopian philosopher, developed rationalist epistemology emphasizing *lebuna* (reason) as supreme authority [23]. Key principles: primacy of reason for discerning truth, skepticism toward authority/dogma, universal rational capacity, ethical rationalism.

XAI Applications: Adversarial Questioning: Don’t accept model outputs uncritically—demand justification through systematic red-teaming and edge case testing. **Democratization:** Universal rationality supports making AI interpretability accessible to all users, not just experts. **Auditing:** Supports third-party auditing, algorithmic

accountability, and transparency mandates.

5 Islamic Logic Traditions

5.1 Kalam (Islamic Theology)

Kalam: Three Evidence Types Kalam (Islamic dialectical theology, 8th-9th centuries CE) [9] distinguishes three knowledge sources: (1) **'Aql**: ratio-nal/intellectual evidence, logical deduction; (2) **Naql**: transmitted/textual evidence, authoritative sources; (3) **Ijma'**: consensus, collective agreement. Robust arguments ideally draw upon all three.

XAI Applications: Hybrid explanation systems mapping to: *'Aql* (model's internal logic, chain-of-thought), *Naql* (training examples, case-based reason-ing), *Ijma'* (ensemble agreement, expert validation). Medical diagnosis exam-ple: "Based on learned correlations (*'aql*), resembling patient ID 4732 (*naql*), confirmed by three models and Dr. Smith (*ijma'*)." When evidence conflicts, Kalam provides adjudication frameworks.

5.2 Falsafa and Sufi Epistemology

Falsafa and Sufi Epistemology **Falsafa** (Islamic philosophy): Ibn Sina (Avi-cenna, 980-1037) and Ibn Rushd (Averroes, 1126-1198) developed modal logic (necessary, possible, impossible) and causal reasoning [24]. **XAI Applications:** Modal distinctions inform uncertainty: necessary predictions (high confidence), possible (medium), impossible (abstain).

Sufism distinguishes *'ilm* (learned knowledge via study) from *ma'rifa* (intu-itive/experiential knowledge, direct gnosis). **XAI Applications:** Not all model behaviors are verbalizable. Deep networks develop "intuitions" resisting expla-nation. Rather than forcing artificial explanations, systems could acknowledge: "This involves patterns too complex to articulate," providing partial explanations while maintaining honesty about interpretability limits.

6 Comparative Analysis

6.1 Structured Comparison

Table 3 compares logic traditions across key dimensions.

Table 3. Comparative Analysis of Logic Traditions for XAI

Tradition	Reasoning Style	Evidence Types	Uncertainty	XAI Strength
Nyaya	Structured inference	4 pramanas	Binary	Systematic explanations
Buddhist	Negative reasoning	Perception + inference	Four-valued	Counterfactuals
Jain	Multi-perspectival	Context-dependent	Seven-fold	Qualified explanations
Yoga	Introspective	Self-examination	Layered	Bias detection
Confucian	Contextual	Pattern + propriety	Situational	Context-aware
Mohist	Empirical	3 standards	Pragmatic	Method evaluation
Ubuntu	Relational	Consensus	Harmony	Relational fairness
Kalam	Dialectical	Multi-source	Hierarchical	Multi-evidence

6.2 Common Themes and Synergies

Recurring themes include multiple forms of knowledge, context dependence, structured argumentation, and relational ontology. Potential synergies include combining Nyaya structure with Ubuntu ethics, or Buddhist negation with Nyaya inference.

6.3 Summary: From Current Practice to Proposed Innovations

Tables 4–6 provide a comprehensive mapping from current XAI methods to proposed non-Western approaches across Indian (Part I), East Asian & African (Part II), and Ethiopian & Islamic traditions (Part III). These serve as the key takeaway, showing *what currently exists*, *what we propose*, and *why it matters*.

Table 4. Summary of Proposed Innovations: Part I – Indian Philosophical Traditions

Tradition	Current XAI Regime	Proposed Change/Addition	Why It Matters
Nyaya	Feature attribution lists (SHAP, IG, LIME) without narrative structure	Five-part structured explanations: Claim -> Evidence -> Universal Principle -> Application -> Conclusion	Pedagogical value; explicit reasoning chains; connects to domain knowledge; culturally legitimate in India
Buddhist Pramana	Binary uncertainty (confident/not confident); post-hoc counterfactuals	(1) Four-valued logic (Is, Is not, Both, Neither) for OOD/ambiguity; (2) Apoha-based counterfactuals (definition by exclusion)	Richer uncertainty quantification; natural handling of paradoxes; aligns with ensemble disagreement
Jain Syadvada	Single-perspective explanations; forced binary classifications	Seven-fold qualified statements acknowledging multiple valid perspectives (syad asti, syad nasti, etc.)	Acknowledges inherent uncertainty; supports multi-stakeholder fairness; allows abstention on indescribable cases
Yoga (Atma-pariksha)	External auditing; post-deployment bias detection	Self-examination frameworks: models analyse own reasoning for biases before answering	Proactive rather than reactive; aligns with Constitutional AI; introspective model design

Table 5. Summary of Proposed Innovations: Part II – East Asian & African Traditions

Tradition	Current XAI Regime	Proposed Change/Addition	Why It Matters
Neo-Confucian (Li-Qi)	Flat explanation treating all features equally	Hierarchical explanations: Principle-level (li: how model generally works) vs. Instance-level (qi: this specific case)	Separates general vs. specific; helps users understand both model behaviour and individual predictions
Japanese Buddhist (Innen)	Isolated feature importance; additive SHAP values	Relational/holistic explanations tracing dependent origination chains through network	Captures feature interactions; shows emergent causation; aligns with graph-based methods
Mohist (San Biao)	Evaluation focuses on accuracy metrics alone	Three-standard evaluation: (1) Theoretical grounding, (2) Empirical validity, (3) Practical utility	Holistic assessment; ensures methods aren't just accurate but theoretically sound and useful
Ubuntu	Individual-level explanations; demographic parity metrics	Relational fairness: community impact analysis; family/network effects; participatory explanation design	Captures systemic harm; contextualises within social relations; community-centred rather than individual

Table 6. Summary of Proposed Innovations: Part III – Ethiopian & Islamic Traditions

Tradition	Current XAI Regime	Proposed Change/Addition	Why It Matters
Ethiopian (ZeraYa-cob)	Trust AI outputs without questioning	Skeptical interrogation protocols: red-teaming, adversarial questioning, explicit justification demands	Promotes critical engagement; democratizes AI literacy; supports accountability and auditing
Kalam ('Aql-Naql-Ijma')	Single explanation source (usually model internals)	Multi-source evidence: (1) Internal logic, (2) Training examples, (3) Expert validation/ensemble consensus	Robustness through triangulation; handles disagreement between methods; hierarchical adjudication
Falsafa (Ibn Sina)	Binary confidence scores	Modal distinctions: Necessary (high confidence), Possible (medium), Impossible (abstain/OOD)	Finer-grained uncertainty communication; aligns with epistemic vs. aleatoric uncertainty
Sufi ('Ilm-Ma'rifa)	Forced verbalization of all decisions	Explicit acknowledgment of inexplicable: some patterns are 'ilm (explainable), others ma'rifa (intuitive, ineffable)	Honesty about interpretability limits; avoids artificial explanations; maintains trust through transparency

Key Insight: These are not replacements but *augmentations*. Current methods (SHAP, LIME, attention) remain valuable. Non-Western traditions offer *additional lenses* that can be deployed contextually depending on audience, domain, and communicative goal.

7 Research Agenda

We outline a research agenda over three time horizons.

7.1 Near-Term Priorities (1–3 Years)

- Nyaya structural probes.** Design probing classifiers that detect *pratijna*, *hetu* and *udaharana* in transformer representations. Use e-SNLI and CoS-E with manual Nyaya-style annotations. We treat this as successful if probe accuracy is above 75% and Sanskrit scholars agree with the labels.
- Apoha-based counterfactuals.** Formalise apoha as an optimisation objective that changes inputs as little as possible to produce a “not-X” case. Compare this with DICE on tabular datasets such as Adult and COMPAS, and on image datasets. We aim for 10–20% smaller perturbations while keeping the counterfactuals valid.
- Ubuntu relational fairness.** Define a notion of relational fairness that measures impact on communities, not only on single individuals. Apply this to recidivism prediction and credit scoring. Show concrete cases where individual fairness and relational fairness give different judgements.

7.2 Medium-Term Goals (3–5 Years)

- Multi-traditional integration.** Build hybrid systems that combine Nyaya-style structure, Buddhist-style uncertainty handling and Ubuntu-style ethics. Test these systems on tasks such as medical diagnosis and loan approval. We see this as successful if more than 70% of users prefer these explanations and fairness violations drop by more than 20%.
- Interdisciplinary validation.** Set up ongoing collaborations with philosophy departments. Present working systems for critique and revision by experts in Nyaya, Buddhism, Jainism and ethics. Signs of success include joint publications and written expert feedback that leads to concrete design changes.

7.3 Long-Term Vision (5+ Years)

6. Unified framework. Develop a meta-framework that can represent several traditions side by side and provide formal mappings between them. Use this to generate explanations that match local cultural norms. In the long run, work with standards bodies such as ISO and IEEE on cross-cultural XAI guidelines, and help bring non-Western logic into AI/ML teaching around the world.

8 Limitations and Future Directions

This survey has clear limits, which we state here.

Scope. We only cover major traditions that have written texts and enough secondary literature. Many oral traditions, local practices and regional ways of knowing are not included because we could not find reliable sources. To fit twelve traditions into one paper, we give an introduction to each rather than a full treatment. We also focus mainly on classical formulations and do not review all modern reinterpretations.

Expertise. We are computer scientists with limited formal training in philosophy. We may have missed important nuances or debates inside each tradition. Many key terms do not have a direct English equivalent, and our translations are approximate. Future work should involve co-authorship with scholars who work directly on Nyaya, Buddhist logic, Jain logic, Ubuntu and Islamic philosophy.

Implementation. Most of our links between logic traditions and XAI are conceptual. Only a few ideas have been implemented as working systems so far. We have not yet run user studies, and many of our mappings are described at an informal or semi-formal level rather than with full mathematics.

Methodology. Our review is broad but not exhaustive. We searched across several databases and libraries, but we did not perform a fully systematic literature review. There is also a deeper tension: we still describe non-Western traditions using existing Western XAI categories. A more radical decolonial approach might question whether these categories should be used at all.

9 Conclusion

Explainable AI does not have to stay within Western intellectual traditions. In this survey we have shown that non-Western logic frameworks — from Nyaya's structured inference to Buddhist apoha, from Ubuntu's focus on relationships to Islamic Kalam's treatment of multi-source evidence — offer rich ideas for interpretability research.

We have offered, to our knowledge, the first cross-tradition survey of non-Western logic traditions aimed at XAI. We identified concrete links to modern ML methods and sketched a research agenda that can be carried out over several years. In the long run, we can imagine XAI systems that generate explanations in more than one epistemic style: Nyaya five-part arguments for users who know Indian logic, Ubuntu-style accounts of community impact for collective decisions, Kalam-style summaries of evidence from different sources, and Buddhist catuṣkoti patterns for nuanced forms of uncertainty. Such systems would not only be more inclusive, they could also be more effective because they draw on a wider range of human ways of knowing.

We invite the research community to take part in this work. ML researchers, philosophers and practitioners who deploy AI in diverse settings all have a role to play. By broadening the conceptual base of XAI, we can design systems that are more culturally grounded, epistemically richer and, in the end, more worthy of trust. This will require real interdisciplinary collaboration, honesty about our own limits and a willingness to engage with perspectives that may at first feel unfamiliar.

Acknowledgments We thank NIELIT for giving us the necessary support to conduct the survey.



References

1. Ribeiro, M.T., Singh, S., Guestrin, C.: "Why Should I Trust You?": Explaining the Predictions of Any Classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 1135–1144 (2016)
2. Lundberg, S.M., Lee, S.-I.: A Unified Approach to Interpreting Model Predictions. In: Advances in Neural Information Processing Systems. pp. 4765–4774 (2017)
3. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention Is All You Need. In: Advances in Neural Information Processing Systems. pp. 5998–6008 (2017)
4. Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E.H., Le, Q.V., Zhou, D.: Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. In: Advances in Neural Information Processing Systems (2022)
5. Toulmin, S.E.: The Uses of Argument. Cambridge University Press, Updated Edition (2003)
6. Matilal, B.K.: The Character of Logic in India. Oxford University Press (1998)
7. Siderits, M., Tillemans, T.J., Chakrabarti, A. (eds.): Apoha: Buddhist Nominalism and Human Cognition. Columbia University Press (2011)
8. Birhane, A.: Algorithmic Injustice: A Relational Ethics Approach. *Patterns* **2**(2), 100205 (2021)
9. Wolfson, H.A.: The Philosophy of the Kalam. Harvard University Press (1976)
10. Makeham, J.: Dao Companion to Neo-Confucian Philosophy. Springer (2010)
11. Priest, G.: The Logic of the Catuskoti. *Comparative Philosophy* **1**(2), 24–54 (2010)
12. Ganeri, J.: Jaina Logic and the Philosophical Basis of Pluralism. *History and Philosophy of Logic* **23**(4), 267–281 (2002)
13. Johnston, I.: The Mozi: A Complete Translation. Columbia University Press (2010)
14. Mohamed, S., Png, M.-T., Isaac, W.: Decolonial AI: Decolonial Theory as So-ciotechnical Foresight in Artificial Intelligence. *Philosophy & Technology* **33**, 659–684 (2020)
15. Basu, A., Chakraborty, T.: Sutra-PCM: A Nyaya-Inspired Neuro-Symbolic Framework for Explainable Commonsense Reasoning. Manuscript under review (2025)
16. Ganeri, J.: Philosophy in Classical India. Routledge (2001)
17. Dreyfus, G.B.: Recognizing Reality: Dharmakirti's Philosophy and Its Tibetan Interpretations. SUNY Press (1997)
18. Wachter, S., Mittelstadt, B., Russell, C.: Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR. *Harvard Journal of Law & Technology* **31**, 841 (2017)
19. Phillips, S.H.: Yoga, Karma, and Rebirth: A Brief History and Philosophy. Columbia University Press (2009)
20. Bai, Y., et al.: Constitutional AI: Harmlessness from AI Feedback. arXiv preprint arXiv:2212.08073 (2022)
21. Madaan, A., Tandon, N., Gupta, P., Hallinan, S., Gao, L., Wiegrefe, S., Alon, U., Dziri, N., Prabhunoye, S., Yang, Y., Gupta, S., Majumder, B.P., Hermann, K., Welleck, S., Yazdanbakhsh, A., Clark, P.: Self-Refine: Iterative Refinement with Self-Feedback. In: Advances in Neural Information Processing Systems (NeurIPS). vol. 36 (2023)
22. Kasulis, T.P.: Zen Action/Zen Person. University of Hawaii Press (1981)
23. Sumner, C.: Ethiopian Philosophy, Vol. II: The Treatise of Zera Yacob and of Walda Heywat: Text and Authorship. Commercial Printing Press, Addis Ababa (1976)
24. Adamson, P.: Philosophy in the Islamic World. Oxford University Press (2016)