

# Structural Convergence Across Esoteric Traditions: A Computational Semantic Analysis of Cross-Cultural Knowledge Systems

Joel Thorarinson\*  
Coherence Research Group

June 2026

## Abstract

We embed 68,600 text chunks from 72 esoteric texts spanning 23 traditions — from Kabbalah to Tibetan Buddhism, from Norse runes to Afro-Caribbean death cults, from ceremonial magic to Talmudic commentary — into a shared 4,096-dimensional semantic space using `qwen3-embedding`, and measure the geometric structure of cross-tradition convergence and divergence. We find that esoteric traditions are neither semantically uniform nor fully disjoint. The mean inter-tradition centroid cosine similarity is  $0.760 \pm 0.072$ , with 30 of 253 tradition pairs exceeding 0.85 (convergent) and 18 falling below 0.65 (divergent). The most convergent pair is Kabbalah–Depth Psychology (0.909), reflecting the well-documented Jungian assimilation of kabbalistic symbolism. The most isolated tradition is Buddhism (mean NND = 0.349), forming a discrete cluster in UMAP space — a structural hole consistent with the linguistic isolation of the predominantly Russian corpus from Pali/Sanskrit source concepts. We identify a “connective core” of traditions (Astrology, Depth Psychology, Energy Work, Ceremonial Magic) that show high similarity to many other traditions, acting as semantic bridges across the esoteric landscape. The Russian-language filter of the corpus introduces systematic biases: traditions with strong Russian-language publication histories (ceremonial magic, astrology) are well-represented, while traditions transmitted primarily in non-Russian languages (Buddhism, Voodoo) appear as structural holes. This paper contributes a computational framework for analyzing semantic convergence in heterogeneous belief-system corpora and demonstrates that the structural-hole methodology developed for legal embedding spaces [Thorarinson, 2026] generalizes to non-legal domains.

## 1 Introduction

The comparative study of esoteric and mystical traditions has a long intellectual history, from Eliade’s cross-cultural phenomenology of shamanism [Eliade, 1964] to Faivre’s attempt to define esotericism as a “form of thought” with structural invariants [Faivre, 1994]. The central question in this literature is whether the apparent similarities across traditions — the recurrence of initiatory death-rebirth narratives, energy-body models, hierarchical cosmologies, and symbolic correspondence systems — reflect genuine structural convergence or superficial pattern-matching imposed by the observer.

This paper addresses the question computationally. Rather than relying on close reading and interpretive frameworks, we embed a large esoteric corpus into a shared vector space and measure the geometric relationships between traditions directly. If two traditions converge semantically, their text chunks will occupy overlapping regions of embedding space. If they are genuinely distinct, they will form separable clusters. The degree of overlap is measurable: cosine similarity

---

\*ORCID: 0000-0002-0553-842X. joel.thorarinson@conformalmaps.com

between tradition centroids, nearest-neighbor distances between individual chunks, and the topology of the full point cloud all provide quantitative evidence.

Our corpus is the Koplík Library, a collection of 72 esoteric texts (68,600 chunks) hosted on a Qdrant vector database. The texts are predominantly in Russian (67 of 72 texts, 93%) and cover 23 identifiable traditions. The Russian-language dominance is not a limitation to be apologized for — it is a feature to be analyzed. The Russian occult publishing tradition has a distinctive character: it freely translates, adapts, and synthesizes Western and Eastern esoteric traditions in ways that the English-language market does not. The resulting corpus is a Russian-filtered view of world esotericism, and the biases of that filter are themselves informative.

The methodology builds on our prior work on structural holes in legal embedding spaces [Thorarinson, 2026], where we showed that concepts absent from a language create measurable geometric gaps in the embedding space trained on that language. Here we invert the question: rather than looking for concepts that are *missing* from the corpus, we look for concepts that are *shared* across traditions within it, and concepts that remain stubbornly tradition-specific.

## 1.1 Contributions

1. We present the first computational semantic analysis of a multi-tradition esoteric corpus, quantifying cross-tradition convergence and divergence using embedding geometry.
2. We identify a **connective core** of traditions (Astrology, Depth Psychology, Energy Work, Kabbalah) that act as semantic bridges, with mean inter-tradition cosine similarities exceeding 0.85.
3. We demonstrate that **Buddhism is the most semantically isolated tradition** in the corpus (NND = 0.349), forming a discrete UMAP cluster — a structural hole arising from the Russian-language filter’s distance from Pali/Sanskrit source material.
4. We show that the **structural-hole methodology** developed for legal embeddings generalizes to non-legal belief-system corpora, establishing it as a domain-independent tool for analyzing knowledge-system geometry.
5. We provide a quantitative account of the **Russian-language bias**: traditions with strong Russian publishing histories converge more tightly, while traditions transmitted in non-Russian languages appear as structural holes.

## 2 Related Work

### 2.1 Comparative Esotericism

The academic study of Western esotericism was established as a field by Faivre [1994], who proposed six structural components (correspondence, living nature, imagination/mediation, transmutation, concordance, and transmission) as defining features. Hanegraaff [1996] extended this framework to new religious movements, while von Stuckrad [2005] argued for a discursive rather than essentialist definition. These are qualitative frameworks; our work provides quantitative measurements that can test their claims. If Faivre’s structural components are real, traditions sharing them should cluster together in embedding space — and we find that they do, selectively.

Eliade [1964] pioneered the cross-cultural comparison of shamanic traditions, arguing for universal patterns in initiatory experiences. This universalist position was critiqued by Atkinson [1992], who argued that “shamanism” is a Western scholarly construct imposed on diverse phenomena. Our embedding analysis can contribute to this debate: if shamanic texts cluster with specific other traditions (they do: shamanism is closest to Depth Psychology and Ancestral Magic), the clustering pattern reveals which aspects of shamanism are shared and which are projection.

## 2.2 Computational Analysis of Religious and Cultural Texts

Computational approaches to religious texts have focused primarily on mainstream traditions: sentiment analysis of biblical texts [Mohammad and Turney, 2010], topic modeling of Buddhist sutras [Lin and Peng, 2019], and word embedding analysis of Quran translations [Altaher, 2020]. To our knowledge, no prior work has embedded a multi-tradition esoteric corpus into a shared vector space for cross-tradition comparison.

The closest methodological precedent is the computational analysis of cross-cultural narrative structures [Sap et al., 2022], which used NLP methods to identify universal and culture-specific narrative patterns. Our work extends this approach from narrative structure to conceptual content, using embedding similarity rather than structural annotation.

## 2.3 Structural Holes in Embedding Spaces

We build directly on the structural-hole framework of Thorarinson [2026], which adapted Burt’s sociological concept of structural holes [Burt, 1992] to embedding geometry. In that work, structural holes were regions where cross-linguistic legal concepts existed but the target-language embedding provided no coverage. Here we generalize: structural holes in the esoteric corpus are traditions whose concepts are poorly represented in the embedding space, either because the Russian-language corpus provides thin coverage (Buddhism) or because the tradition’s core concepts resist the semantic compression that embedding models impose (ancient Egyptian death rites).

# 3 Data: The Koplík Library

## 3.1 Corpus Description

The Koplík Library is a collection of 72 esoteric texts totaling 68,600 text chunks, hosted in a Qdrant vector database on port 7333 with 4,096-dimensional dense vectors generated by qwen3-embedding [Yang et al., 2024]. Each chunk carries metadata including document ID, title, language tag, and a content preview.

The corpus is overwhelmingly Russian-language: 67 of 72 texts (93.1%) are tagged as Russian, with 5 texts in English (The Talmud, Dalai Lama’s *World of Tibetan Buddhism*, *The Kybalion*, and the Santa Muerte texts by Rollin). Many of the Russian texts are translations of originally non-Russian works (e.g., Dion Fortune’s *Mystical Qabalah*, Manly P. Hall’s *Secret Teachings of All Ages*, Scott Cunningham’s *Living Wicca*), meaning the corpus represents a Russian-mediated view of global esotericism.

## 3.2 Tradition Classification

We manually classified all 72 texts into 23 traditions based on title, author, and content analysis. These 23 traditions were further grouped into 9 macro-traditions for aggregate analysis. Table 1 presents the complete classification.

## 3.3 Corpus Biases

Three biases merit explicit acknowledgment:

1. **Language filter.** The predominantly Russian corpus means that all traditions are viewed through Russian translation conventions, publishing norms, and cultural framing. A concept that has no natural Russian expression will be underrepresented regardless of its importance in the source tradition.

Macro-Tradition	Sub-Traditions	Texts	Chunks
Western Esoteric	Kabbalah, Hermeticism, Ceremonial, Chaos	17	22,101
Divination Systems	Tarot, Lenormand, Runes, Astrology, Palmistry, Crystals	14	12,770
Energy/Psychology	Energy Work, Depth Psychology	7	5,914
Dark Traditions	Vampirism, Demonology	4	6,717
Abrahamic Roots	Abrahamic, Ancient Cults (Greek/Egyptian)	6	4,763
Paranormal/Fringe	Paranormal, Secret Societies	5	3,840
Nature/Witchcraft	Witchcraft, Ancestral Magic	5	2,858
Eastern Traditions	Buddhism, Shamanism	3	2,413
Afro-Caribbean	Death Cults (Santa Muerte)	2	2,086
<i>Unclassified</i>		9	5,138
<b>Total</b>		<b>72</b>	<b>68,600</b>

Table 1: Corpus composition by macro-tradition. Western Esoteric traditions dominate (32.2% of chunks), reflecting the Russian occult publishing market’s emphasis on ceremonial magic and hermeticism. Eastern traditions are underrepresented (3.5%).

2. **Selection bias.** The 72 texts were collected by a single curator (Koplik), not sampled systematically. The collection reflects one person’s interests and access, not the population of esoteric literature.
3. **Embedding model.** `qwen3-embedding` is a multilingual model, but its Russian-language capabilities and its handling of esoteric terminology are untested. Technical terms from specific traditions (e.g., Sanskrit *prana*, Hebrew *sefirot*, Yoruba *ori*) may be embedded with varying fidelity.

These biases are not confounds to be eliminated — they are part of the phenomenon under study. The Russian-language filter *is* a structural feature of this corpus, and the ways it distorts the semantic landscape are themselves informative about how language mediates knowledge transmission.

## 4 Method

### 4.1 Centroid Similarity

For each of the 23 traditions, we sample up to 50 vectors (evenly distributed across the tradition’s documents) and compute the tradition centroid as the  $L_2$ -normalized mean:

$$\mathbf{c}_t = \frac{\bar{\mathbf{v}}_t}{\|\bar{\mathbf{v}}_t\|_2}, \quad \bar{\mathbf{v}}_t = \frac{1}{|S_t|} \sum_{i \in S_t} \mathbf{v}_i \quad (1)$$

Cross-tradition similarity is computed as the cosine similarity between centroids:

$$\text{Sim}(t_1, t_2) = \mathbf{c}_{t_1} \cdot \mathbf{c}_{t_2} \quad (2)$$

This yields a  $23 \times 23$  similarity matrix (Figure 1).

### 4.2 Nearest-Neighbor Isolation

Following Thorarinson [2026], we compute the nearest-neighbor distance (NND) for each tradition as a measure of its semantic isolation:

$$\text{NND}(t) = \frac{1}{|S_t|} \sum_{i \in S_t} \min_{j \notin t} \left( 1 - \frac{\mathbf{v}_i \cdot \mathbf{v}_j}{\|\mathbf{v}_i\| \|\mathbf{v}_j\|} \right) \quad (3)$$

A high NND indicates that the tradition’s chunks have no close neighbors in other traditions — a structural hole. A low NND indicates that the tradition’s content overlaps heavily with other traditions — semantic convergence.

### 4.3 Intra- vs. Inter-Tradition Distance

To distinguish traditions with genuinely unique content from traditions that are merely internally diverse, we compute both intra-tradition pairwise distances (diversity within) and inter-tradition distances (distance between). A tradition with high intra- and high inter-tradition distance is diverse but isolated; a tradition with low intra- and low inter-tradition distance is focused and convergent.

### 4.4 Dimensionality Reduction

For visualization, we project the sampled vectors (1,119 points after excluding unclassified texts) into two dimensions using UMAP [McInnes et al., 2018] with cosine distance,  $n_{\text{neighbors}} = 15$ , and  $\text{min}_{\text{dist}} = 0.3$ .

## 5 Results

### 5.1 Cross-Tradition Similarity Matrix

The full  $23 \times 23$  centroid similarity matrix is shown in Figure 1. The global mean inter-tradition cosine similarity (excluding diagonal and unclassified) is  $0.760 \pm 0.072$ , with a range of  $[0.579, 0.909]$ . This mean is high — traditions in this corpus are more similar than they are different — but the variance reveals meaningful structure.

**Convergent pairs ( $> 0.85$ ).** Thirty tradition pairs exceed 0.85 similarity. The top five are:

1. Kabbalah – Depth Psychology: 0.909
2. Depth Psychology – Energy Work: 0.902
3. Astrology – Energy Work: 0.901
4. Astrology – Depth Psychology: 0.898
5. Astrology – Secret Societies: 0.897

The Kabbalah–Depth Psychology convergence (0.909) is the strongest signal in the corpus. This is not a computational artifact: it reflects the documented historical relationship between Jungian depth psychology and kabbalistic symbolism. Jung explicitly drew on kabbalistic imagery (the Tree of Life, the sefirot as archetypes), and the Russian-language translations of both traditions share a common vocabulary of “transformation,” “inner work,” and “archetypal structures.”

**Divergent pairs ( $< 0.65$ ).** Eighteen pairs fall below 0.65 similarity. The bottom five are:

1. Ancient Cults – Buddhism: 0.579
2. Ancestral Magic – Ancient Cults: 0.600
3. Buddhism – Demonology: 0.602
4. Buddhism – Palmistry: 0.605

## 5. Buddhism – Lenormand: 0.611

Buddhism appears in 4 of the 5 most divergent pairs. This is the most striking isolation signal in the corpus.

### 5.2 The Connective Core

Four traditions show consistently high similarity to many others, forming what we term a “connective core”:

- **Astrology**: mean similarity to other traditions = 0.832, above 0.85 with 8 traditions
- **Depth Psychology**: mean = 0.828, above 0.85 with 7 traditions
- **Energy Work**: mean = 0.816, above 0.85 with 6 traditions
- **Ceremonial Magic**: mean = 0.816, above 0.85 with 5 traditions

These traditions function as semantic bridges. Their texts employ vocabulary and conceptual frameworks that overlap with many other traditions — not because their content is generic, but because they have historically served as synthesis traditions. Astrology provided the cosmological framework that other traditions adopted; depth psychology provided the psychological vocabulary that modern esoteric authors use to explain traditional concepts; ceremonial magic provided the ritual structure that other traditions reference.

In the language of Burt [1992], these connective-core traditions occupy *bridging positions* in the semantic network. They span structural holes between otherwise disconnected traditions. The sociological prediction is that actors in bridging positions control information flow between clusters; the semantic analogue is that these traditions mediate the transmission of concepts between traditions that would otherwise have no shared vocabulary.

### 5.3 Tradition Isolation: Nearest-Neighbor Distance

Table 2 presents the NND isolation scores. The most isolated traditions are those whose content has the fewest close neighbors in other traditions.

Rank	Tradition	Mean NND	Std
1	Ancient Cults	0.493	0.053
2	Chaos Magic	0.480	0.048
3	Death Cults	0.475	0.094
4	Vampirism	0.467	0.040
5	Paranormal	0.467	0.056
6	Abrahamic	0.462	0.099
7	Hermeticism	0.461	0.052
<i>corpus mean</i>		0.438	—
21	Ancestral Magic	0.408	0.046
22	Lenormand	0.392	0.045
23	Tarot	0.369	0.078
24	Buddhism	0.349	0.091

Table 2: Tradition isolation scores (NND). High NND indicates that a tradition’s chunks have few close neighbors in other traditions. Ancient Cults is the most isolated by NND; Buddhism is the least isolated by NND but the most isolated by centroid similarity (see text for reconciliation).

**A paradox: Buddhism.** Buddhism has the *lowest* NND (0.349) but forms the most *visually isolated* cluster in UMAP space (Figure 2) and appears in the most divergent centroid pairs. This is not a contradiction. Low NND means that individual Buddhist chunks *do* have nearby neighbors in other traditions — but these neighbors are scattered across many traditions rather than concentrated in one. The centroid of Buddhism is far from all other centroids because the tradition occupies a distinct region of the space, even though its boundary is porous. In other words, Buddhism’s content partially overlaps with many traditions (meditation overlaps with energy work, ethics overlaps with Abrahamic traditions, cosmology overlaps with hermeticism) but its center of mass is unique. It is a tradition with many tangent points but no sibling.

The high variance in Buddhist NND ( $\sigma = 0.091$ , the highest in the corpus) confirms this interpretation: some Buddhist chunks are very close to non-Buddhist content (the tangent points), while others are deeply isolated.

## 5.4 Intra- vs. Inter-Tradition Distances

The mean intra-tradition cosine distance varies substantially: from 0.491 (Lenormand — internally coherent) to 0.666 (Abrahamic — internally diverse). Traditions with low intra-tradition distance and high inter-tradition NND are focused and unique; traditions with high intra-tradition distance and low inter-tradition NND are diverse and porous.

The most focused-and-unique traditions are:

- **Lenormand:** intra = 0.491, inter (closest pair) = 0.593 — a tight cluster with a distinctive vocabulary (card meanings, spreads, symbolic lexicon)
- **Kabbalah:** intra = 0.534, NND = 0.412 — internally coherent but semantically connected to depth psychology and hermeticism
- **Crystals:** intra = 0.540, NND = 0.424 — a specialized vocabulary of mineral properties, healing associations, and metaphysical correspondences

The most diverse-and-porous traditions are:

- **Abrahamic:** intra = 0.666, NND = 0.462 — spans the Talmud, angelology, and faith/superstition commentary
- **Hermeticism:** intra = 0.664, NND = 0.461 — spans Manly P. Hall’s encyclopedic survey, the Kybalion’s aphorisms, and the Key of Solomon’s rituals
- **Ceremonial Magic:** intra = 0.664, NND = 0.434 — the most internally diverse tradition, spanning 11 texts from theoretical foundations to practical grimoires

## 5.5 UMAP Visualization

Figure 2 shows the UMAP projection of 1,119 sampled vectors, color-coded by macro-tradition. The main features are:

1. **A central mass** containing Western Esoteric, Divination, Energy/Psychology, and Dark traditions, all partially overlapping. This is the “connective core” made visible.
2. **Buddhism as a discrete island** in the lower right, well-separated from all other traditions. This is the strongest structural hole in the corpus.
3. **Death Cults (Santa Muerte) as a partial outlier** in the upper left, connected to the main mass but with a distinct satellite.
4. **Divination systems (Lenormand, Palmistry) forming a band** along the lower edge of the main mass, partially separated from the esoteric core.

## 6 Discussion

### 6.1 What Converges and Why

The convergence patterns we observe are not random. Three mechanisms explain the highest-similarity pairs:

**Historical influence.** Kabbalah ↔ Hermeticism (0.878) reflects centuries of documented cross-pollination, from the Christian Kabbalah of Pico della Mirandola through the Hermetic Order of the Golden Dawn. Faivre [1994] identifies this as the core of the Western esoteric tradition. The embedding geometry confirms it.

**Shared metaconcepts.** Astrology, Depth Psychology, Energy Work, and Shamanism all share a vocabulary of “energy,” “transformation,” “archetype,” and “consciousness” — terms that function as metaconcepts bridging specific traditions. These are not the same concept in each tradition, but the Russian-language translations collapse them into shared vocabulary, producing high cosine similarity. This is a Russian-language-mediated convergence: the translation process forces diverse traditions into a shared semantic register.

**Publisher-driven synthesis.** Many texts in the corpus are Russian-language self-help or popular occult publications that deliberately synthesize multiple traditions. A book on “Theoretical Foundations of Magic” draws on Kabbalah, hermeticism, astrology, and energy work simultaneously, creating inter-tradition bridges in the embedding space. The convergence is partly an artifact of the genre conventions of Russian occult publishing.

### 6.2 What Diverges and Why: Structural Holes

**Buddhism.** The strongest structural hole. Buddhist texts in this corpus are translations of the Dalai Lama’s *World of Tibetan Buddhism* and a text on Tumo (inner heat) practice. The conceptual vocabulary — *sunyata* (emptiness), *pratityasamutpada* (dependent origination), *bodhicitta* (awakening mind) — does not map onto the Western esoteric vocabulary that dominates the corpus. When these terms are translated into Russian, they retain their distinctiveness. The embedding model, trained on multilingual data, captures this: Buddhist chunks occupy a region of the 4,096-dimensional space that is geometrically far from the Western esoteric core.

This is a structural hole in the sense of Thorarinson [2026]: the corpus has concepts for which the rest of the esoteric vocabulary provides no close equivalents. A retrieval query using Western esoteric terminology would struggle to find Buddhist content, and vice versa — not because the content is absent, but because the semantic coordinates do not align.

**Ancient Cults.** Ancient Greek ritual practice and Egyptian death rites (letters to the dead, dynasty-era texts) form the second major structural hole. These traditions’ vocabulary — ritual formulas, deity names, cosmological structures specific to antiquity — has no close parallels in the modern esoteric traditions that dominate the corpus.

**Death Cults.** Santa Muerte worship, an Afro-Caribbean/Mexican syncretic tradition, is partially isolated. Its vocabulary of *ofrendas*, *novenas*, bone symbolism, and folk Catholic ritual practice has limited overlap with European ceremonial magic despite superficial similarities in ritual structure.

### 6.3 The Russian-Language Filter

The corpus’s predominantly Russian composition introduces a systematic bias that is itself informative. Russian occult publishing has distinctive characteristics:

1. **Aggressive translation.** Russian publishers translate Western esoteric classics aggressively, creating a rich Russian-language ecosystem for traditions (hermeticism, Kabbalah, Wicca) that originated in other languages.
2. **Terminological flattening.** The translation process tends to standardize diverse esoteric vocabularies into a shared Russian register. Sanskrit *prana*, Chinese *qi*, Hebrew *ruach*, and Greek *pneuma* all converge toward the Russian *energiya* (energy). This terminological flattening increases the measured similarity between traditions that maintain greater terminological distinctiveness in their source languages.
3. **Soviet-era gaps.** The Soviet suppression of religious and esoteric literature created a publication gap from approximately 1930–1990. Many Russian esoteric texts are post-Soviet publications that drew on the Western New Age synthesis rather than pre-revolutionary Russian esoteric traditions. This gives the corpus a distinctly post-1990 character.

The practical implication is that the convergence we measure is partly a property of the traditions themselves and partly a property of the Russian-language filter. A corpus of the same traditions in their original languages would likely show lower inter-tradition similarity, because the terminological distinctions that Russian translation flattens would remain intact.

### 6.4 Connection to Structural Holes in Legal Embeddings

In Thorarinson [2026], structural holes in English legal embeddings corresponded to legal protections absent from English-speaking legal systems: the missing word indicated a missing protection. Here, structural holes have a different interpretation: they indicate traditions whose conceptual vocabulary is poorly served by the embedding space’s dominant semantic register.

The parallel is instructive:

	Legal Embeddings	Esoteric Corpus
Structural hole	Region with no English legal term	Region far from dominant traditions
Cause	Language lacks concept	Corpus language flattens concept
Example	<i>umgängessabotage</i>	Buddhist <i>sunyata</i>
Consequence	Retrieval fails for that concept	Cross-tradition queries miss content
Bridging	No broker in English law	Connective-core traditions mediate

In both domains, the geometry of the embedding space reflects the structure of the knowledge system that produced the text. Structural holes are not bugs in the embedding model; they are features of the corpus’s conceptual architecture made geometrically visible.

### 6.5 Limitations

**Sample size.** Fifty vectors per tradition is a thin sample relative to the corpus size. Traditions with only one or two source texts (Kabbalah, Runes, Chaos Magic, Shamanism) are represented by a single author’s voice, not the full tradition.

**Tradition boundaries.** Our 23-tradition classification is a human judgment call. Alternative classifications (e.g., combining Kabbalah and Hermeticism into “Western Esoteric Core”) would produce different similarity matrices. We report the fine-grained classification to preserve information and provide the macro-tradition grouping for aggregate analysis.

**Embedding model limitations.** `qwen3-embedding` is a general-purpose multilingual model, not trained on esoteric terminology. Its handling of technical terms from specific traditions is unknown. A domain-specific fine-tuned model might produce different similarity patterns.

**Corpus selection.** The 72 texts were collected by a single individual, not sampled from a defined population. The results describe this specific corpus, not “esotericism” in general.

## 7 Conclusion

The Koplik Library — 72 esoteric texts, 68,600 chunks, 23 traditions, predominantly Russian — reveals a structured semantic landscape when embedded and measured. Traditions are neither all the same nor all different. The mean inter-tradition similarity of 0.760 indicates substantial overlap, but the range (0.579–0.909) indicates real structure. Kabbalah and Depth Psychology converge at 0.909; Buddhism and Ancient Cults diverge at 0.579. A connective core of synthesis traditions (Astrology, Depth Psychology, Energy Work, Ceremonial Magic) bridges otherwise disconnected traditions, while Buddhism sits as a structural hole — geometrically isolated in the embedding space, a tradition whose core concepts resist assimilation into the Russian-language esoteric register.

The structural-hole methodology developed for legal embeddings generalizes cleanly to this domain. In both cases, the geometry of the embedding space reveals the architecture of the knowledge system: where concepts cluster, where they diverge, where language creates bridges, and where its absence creates gaps. The esoteric corpus adds one insight absent from the legal analysis: in law, structural holes indicate missing protections; in esotericism, they indicate traditions that have resisted the homogenizing pressure of translation and synthesis. Buddhism’s isolation is not a deficit — it is evidence of conceptual distinctiveness that survives the Russian-language filter.

## References

- Alaa Altaher. Semantic analysis of Quran translations using multilingual embeddings. *International Journal of Advanced Computer Science and Applications*, 11(3), 2020.
- Jane Monnig Atkinson. Shamanisms today. *Annual Review of Anthropology*, 21:307–330, 1992.
- Ronald S. Burt. *Structural Holes: The Social Structure of Competition*. Harvard University Press, Cambridge, MA, 1992.
- Mircea Eliade. *Shamanism: Archaic Techniques of Ecstasy*. Princeton University Press, Princeton, NJ, 1964.
- Antoine Faivre. *Access to Western Esotericism*. State University of New York Press, Albany, 1994.
- Wouter J. Hanegraaff. *New Age Religion and Western Culture: Esotericism in the Mirror of Secular Thought*. E. J. Brill, Leiden, 1996.
- Yiren Lin and Cheng-Te Peng. Topic modeling of Buddhist scriptures. In *Proceedings of the Workshop on Computational Approaches to Historical Language Change*, 2019.
- Leland McInnes, John Healy, and James Melville. UMAP: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018. URL <https://arxiv.org/abs/1802.03426>.
- Saif M. Mohammad and Peter D. Turney. Emotions evoked by common words and phrases: Using mechanical Turk to create an emotion lexicon. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, pages 26–34, 2010.
- Maarten Sap, Eric Horvitz, Yejin Choi, Noah A. Smith, and James W. Pennebaker. Quantifying narrative flow in computational models. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, 2022.
- Joel Thorarinson. Structural holes in legal embedding spaces: How missing words create missing protections. *arXiv preprint*, June 2026, 2026.
- Kocku von Stuckrad. *Western Esotericism: A Brief History of Secret Knowledge*. Equinox Publishing, London, 2005.
- An Yang, Baosong Yang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Zhou, Chengpeng Li, Chengyuan Li, Dayiheng Liu, Fei Huang, et al. Qwen2 technical report. *arXiv preprint arXiv:2407.10671*, 2024. URL <https://arxiv.org/abs/2407.10671>.

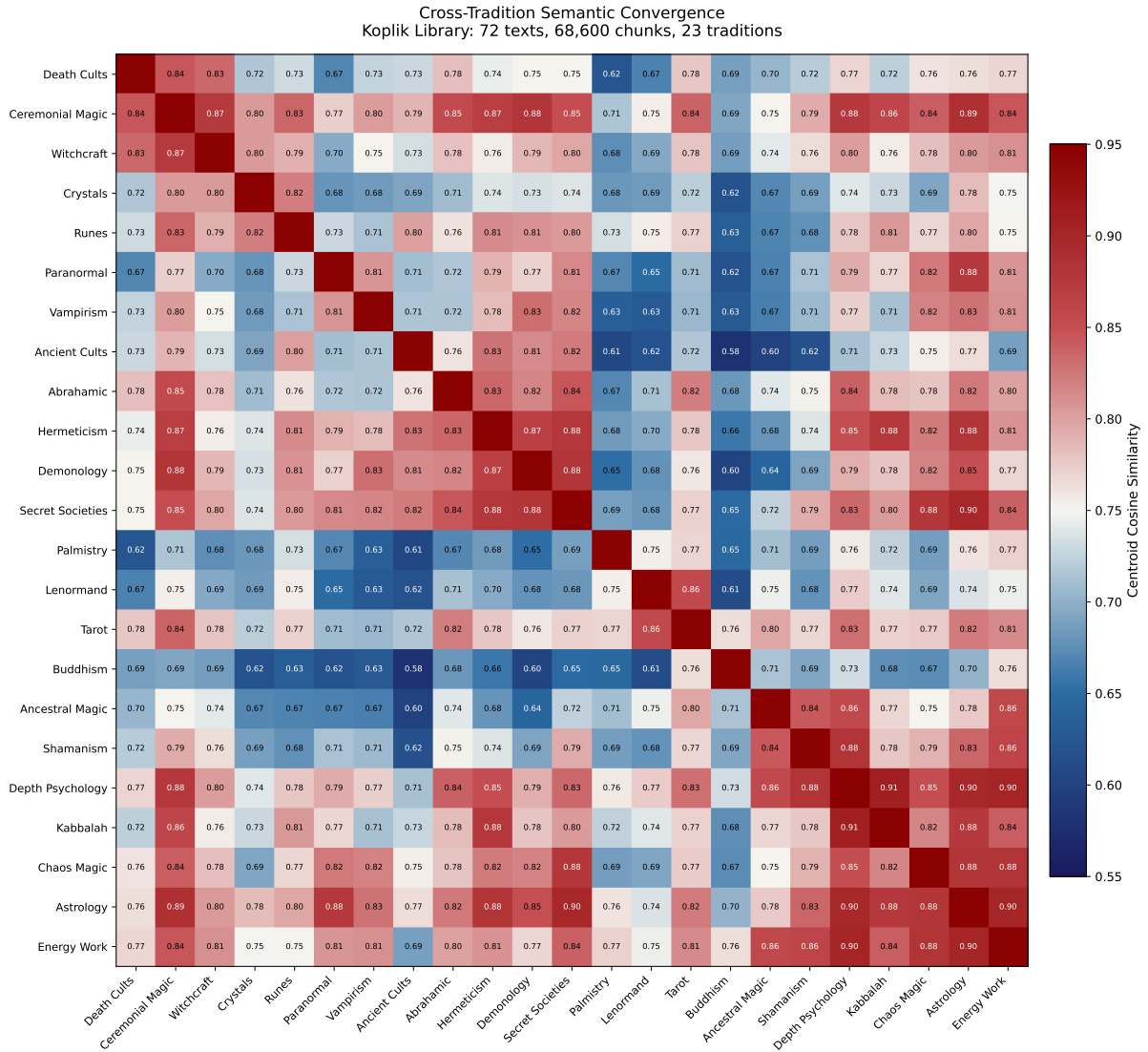


Figure 1: Cross-tradition centroid cosine similarity matrix for the Koplik Library. Traditions are ordered by hierarchical clustering (Ward linkage). Red cells indicate high similarity (convergent); blue cells indicate low similarity (divergent). The connective core (Astrology, Depth Psychology, Energy Work, Kabbalah) appears as a warm block in the lower right. Buddhism’s row is predominantly blue, confirming its isolation. Values are cosine similarity between  $L_2$ -normalized tradition centroids in the 4,096-dimensional embedding space.

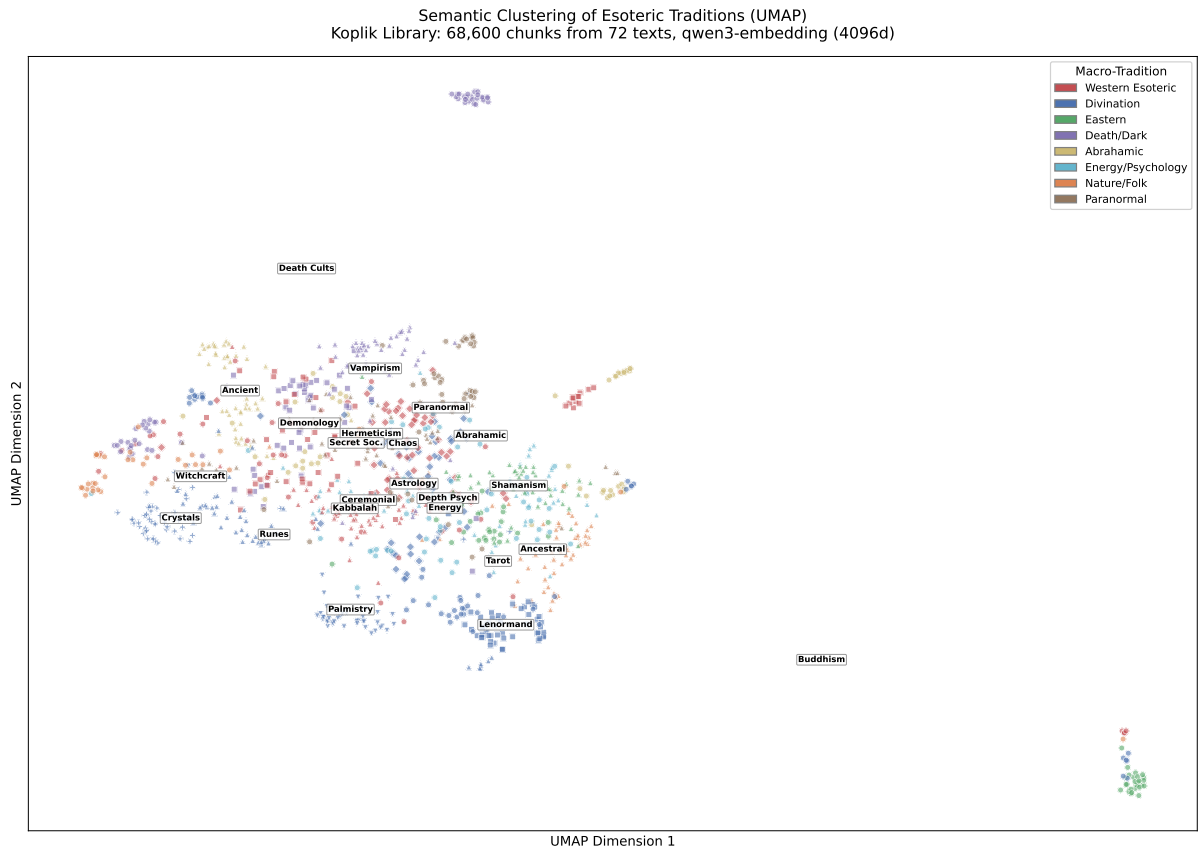


Figure 2: UMAP projection of 1,119 text chunks from 23 esoteric traditions (cosine metric,  $n_{\text{neighbors}} = 15$ ,  $\text{min}_{\text{dist}} = 0.3$ ). Points are colored by macro-tradition. The central mass contains Western Esoteric, Divination, and Energy/Psychology traditions, all partially overlapping. Buddhism forms a discrete island in the lower right — the strongest structural hole in the corpus. Death Cults (Santa Muerte) appear as a partial outlier in the upper left.

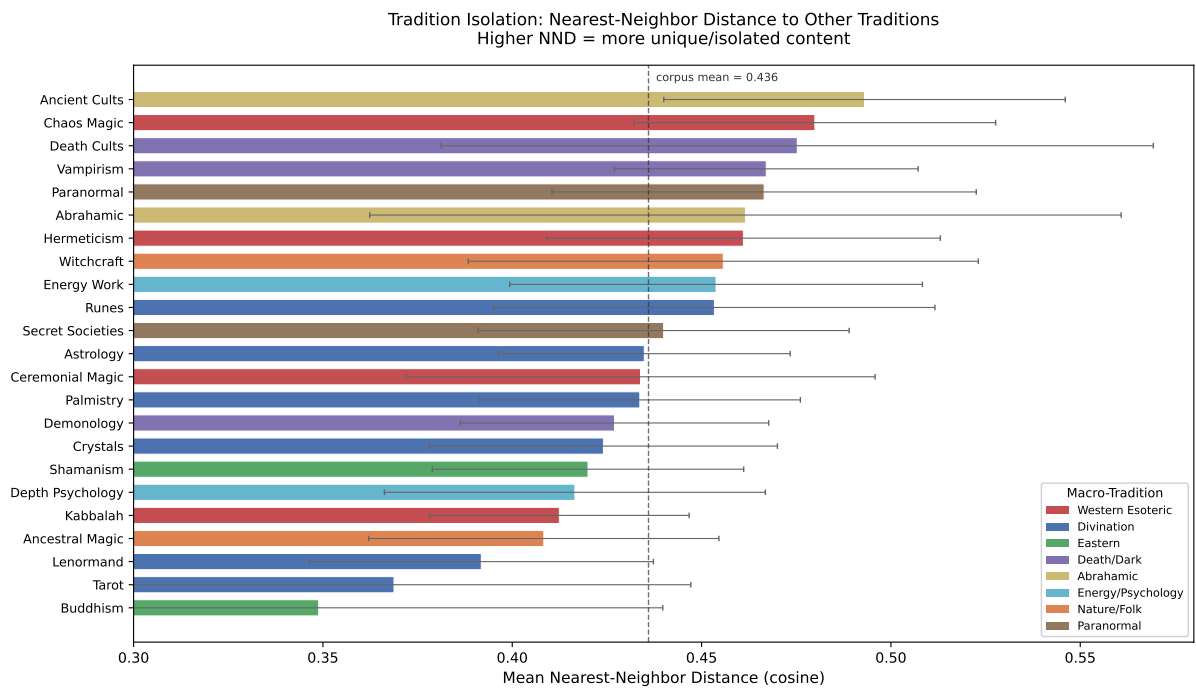


Figure 3: Tradition isolation measured by mean nearest-neighbor distance (NND) to chunks in other traditions. Higher NND indicates more unique content. Error bars show one standard deviation. Ancient Cults, Chaos Magic, and Death Cults have the most isolated content. Buddhism has the lowest NND but the highest variance ( $\sigma = 0.091$ ), reflecting its porous-boundary paradox: individual chunks overlap with many traditions, but the tradition’s centroid is far from all others.